MULTIHOP PACKET RADIO NETWORKS: DESIGN ALGORITHMS AND PROTOCOLS

HUNG, Kwok-Wah

A DISSERTATION SUBMITTED IN PARTIAL
FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY
IN
THE DIVISION OF ELECTRONIC ENGINEERING

GRADUATE SCHOOL

THE CHINESE UNIVERSITY OF HONG KONG

HONG KONG

APRIL 1991

Glory to God

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my research supervisor, Dr. Tak-Shing Yum, for his patient guidence, valuable suggestion and continuous encouragement during the years of research. He shares with me not only up-to-date technology and bright ideas but also with devoted research attitude and friendship.

It is my pleasure also to thank the Department of Computing, Hong Kong Polytechnic, for granting me a full-time study leave to complete my thesis. Special thanks are due to Mr. Vir Gulati, Ms. Janice Burn and Mr. Chris Walsh for their support on my application for the Staff Development Programme.

MULTIHOP PACKET RADIO NETWORKS: DESIGN ALGORITHMS AND PROTOCOLS

ABSTRACT

In this thesis, several design algorithms and communication protocols for multihop Packet Radio Networks (PRNs) are proposed and evaluated. We first give an overview of PRNs and highlight some design issues. We then design the Simple Tone Sense (STS) protocol for multihop Packet Radio Networks with multiple directional antennas stations. The protocol can minimize transmission interference by using a group of tones to identify the active neighbours. A variation of the STS protocol called the Variable Power Tone Sense (VPTS) protocol is also designed to further reduce interference. Algorithms for assigning tones and for determining the orientation and broadcasting angles of the directional antennas are designed. Simulation results show that the STS and VPTS protocols performs particularly well when the traffic is heavy. We then turn to investagate the design issues related to Spread Spectrum Packet Radio Networks (SS/PRNs). We design a spreading code assignment algorithm which could reduce the number of codes required to about 20% of the number of stations in the network. Further reducing the number of codes is found to cause little throughput degradation. The Coded Tone Sense protocol is designed for using these codes in multihop SS/PRNs. We then relate the code assignment problem to the graph coloring problem and propose a very efficient algorithm for assigning codes to the stations. A very tight lower bound on the number of codes needed is also derived. We then turn to design a new scheduling algorithm for packet transmissions. The design objective is to have a schedule with minimum cycle length, maximum network throughput and fair allocation of transmission capacities among all stations. For comparison with other scheduling algorithms in the literatures the following performance measures for scheduling algorithms are derived: (1) the cycle length, (2) the scheduling delay, (3) the minimum transmission capacity and (4) the normalized network capacity. The new algorithm is found to give schedules that have (1) the shortest cycle length, (2) the smallest scheduling delay, (3) the largest minimum transmission capacity and (4) the same highest normalized network capacity when compared to two of the best scheduling algorithms in the literature. Finally we apply the new scheduling algorithm to design the Staggered Multicast Protocol which is suitable for unicasting, broadcasting as well as multicasting in SS/PRNs. The Common-Header/Transmitter-Based spreading code is used for data packets transmission and the receiver-based code is used for acknowledgement packets transmission. By staggering packet transmission the protocol can significantly reduce broadcasting delay. Special addressing method and packet format are also designed to achieve collision-free acknowledgement and multicast capability. Simulation results show that the protocol provides better throughput-delay performance than the Common-Header/Transmitter-Based Slotted ALOHA protocol despite its added capabilities of staggered relay broadcasting, collision-free acknowledgement and global packet multicast.

多段無線電訊包網絡之設計算法和協議

論文提要

本論文提出及評價數項『多段無線電訊包網絡』的設計算法和通訊協議。首先我們概覽了無線電訊包網絡的操作和設計要點。然後我們為置有單向天線的網絡設計了『簡單音調檢測協議』。此協議利用一組音調來識別相鄰站的繁忙情況，從而減低通訊時的干擾。將此協議稍為修改，便設計出『可變功率音調檢測協議』來進一步減低干擾。我們亦設計了相關算法來分配音調、決定天線的傳送方向和廣播角度。模擬結果顯示以上兩種新協議在業務繁忙時表現特佳。我們轉而探討『擴頻無線電訊包網絡』的設計事宜。首先我們設計了一套擴展碼分配算法，將碼的數目減至站的總數的兩成。我們發現進一步減少碼的數目，只會輕微降低通量。『編碼音調檢測協議』便是專門為多段擴頻網絡而設計。我們接著將擴展碼的分配問題和圖的著色問題相連起來，並提出一項很有效的算法來分配擴展碼。我們亦推算了所需碼數量的下限。此外，我們又設計了一種新的訊包傳送規劃算法。目的是使定下的規程能夠使用最短週期，提供最大網絡通量和平均分配傳送容量給各站。為了和文獻中其他的規劃算法比較，我們推算了以下的性能尺度：（1）週期長度，（2）規劃時延，（3）最低傳送容量，（4）標準化網絡容量。當和文獻中兩種優良的規劃算法比較時，我們設計的新算法有以下優點：（1）週期最短，（2）規劃時延最短，（3）最低傳送容量最大，（4）相近的高標準化網絡容量。最後我們應用此新算法為多段擴頻網絡設計了『堆疊多址協議』。此協議使用公用訊包頭碼和傳送站為本的擴展碼來傳送數據訊包，而接收站為本的擴展碼則用來傳送應答訊包。由於訊包傳送時能互相堆疊，故廣播時延能顯著地縮短了。此外，我們設計了特別的編址方法和訊包格式以避免應答訊包互撞和達到多址功能。

Published and Submitted Papers from This Thesis

[1] "The Coded Tone Sense Protocol for Multi-Hop Spread Spectrum Packet Radio Networks", Proc. of IEEE GLOBECOM'89, Dallas, U.S.A. November 1989.

[2] "Design Algorithms for Multi-Hop Packet Radio Networks with Multiple Directional Antennas stations", Proc. of IEEE GLOBECOM'89, Dallas, U.S.A. November 1989; also accepted for publication in IEEE Transactions on Communications, to be appeared in 1991.

[3] "An Efficient Code Assignment Algorithm for Multi-Hop Spread Spectrum Packet Radio Networks", Proc. of IEEE GLOBECOM'90, San Diego, U.S.A. December 1990.

[4] "Fair and Efficient Transmission Scheduling in Multihop Packet Radio Networks", submitted to IEEE Transactions on Communications for possible publication.

[5] "Staggered Multicast Protocol with Collision-free Acknowledgement in Multihop Spread Spectrum Packet Radio Networks", submitted to IEEE GLOBECOM'91 for possible publication.

TABLE OF CONTENTS

CHAPTER 1

OVERVIEW OF PACKET RADIO NETWORKS

1.1 INTRODUCTION

Packet Radio Networks (PRNs) apply packet switching technology to communicate among stations via a broadcast radio medium. They are useful for communication in regions where wire connection between users is not practical or expensive. This wireless network is particular suitable for communication among mobile users. Besides the connection feature the main difference between point-to-point networks and PRNs is that the radio channel is both a multiaccess and a broadcast medium. Therefore the existing techniques used in point-to-point networks might not be readily applicable to PRNs.

The ALOHA system at the University of Hawaii [ABRA 70] is the first computer system employing radio transmissions. As the interest in PRNs grew, many research and development effort have been spent in designing such networks. The design issues include signalling method, network topology, channel access protocol, data link control, routing algorithm, network management and many others [LEIN 87].

When the size of a network is small, the transmission power of each station can cover all other stations and so the network is fully connected. As the network size gets larger, a multihop network involving packet relaying is usually more suitable for connecting all stations. In multihop PRNs many approaches to topological design and associated channel access protocols can be used to enhance system performance by spatial reuse of the communication channel. In spread spectrum communications the use of spreading codes permits a receiver to extract a particular signal from many overlapping ones and adds another dimension in the design of PRNs.

In this chapter we first introduce the structure of a PRN and its components. We then discuss briefly the design issues relating to channel access protocols, spatial reuse effect and spread spectrum usage. Detailed and excellent discussion on these matters can be found in the special issue on PRNs, Proceeding of IEEE in January 1987.

1.2 NETWORK STRUCTURE

In PRNs each station is equipped with a radio unit, an antenna and a digital controller. The radio unit and the antenna provide the connectivity between neighbouring stations. When two stations are within the transmission range of each other they are able to communicate directly. When a station wants to send a packet to a more distant station, the packet is relayed based on a store-and-forward operation. Omnidirectional antenna is usually used for transmission but directional antenna can also be used to reduce transmission interference. If there is overlap of transmissions from different stations, all the packets involved would be destroyed.

The digital controller provides the packet switching function for relaying packets to their final destinations. It receives packets from neighbouring stations, makes routing decisions and forwards packets on to the next station in the routing path. The controller also provides other network management functions such as flow and congestion control.

When the network does not achieve sufficient connectivity, repeaters are to be placed in appropriate locations to provide the necessary connection. A repeater has the same packet switching function as a normal station except ~~expect~~ that the repeater itself will not be a source or a final destination. Usually there is a central control station responsible for gathering connectivity information, computing the routing table and distributing the routing information. Various network control functions are also provided by this control station. In some PRNs multiple control stations are used while others may adopt a fully distributed control algorithm.

1.3 Channel Access Protocol

In sharing the multiaccess radio channel, suitable method should be used to resolve the contention among neighbouring stations. Time division and code division techniques can be used in spread spectrum transmissions and this is to be discussed in the following sections. When all stations use the same frequency for narrow-band transmissions, fixed TDMA can be used to assign time slots to stations so as to prevent collision. However this method gives low channel utilization when the data traffic is bursty. Random-access protocols are more suitable for bursty traffic especially when the network requires short transmission delay. There are many channel access protocols found in the literature [TOBA 80]. The Slotted ALOHA protocol, the Carrier Sense Multiple Access (CSMA) protocol and the Busy Tone Multiple Access (BTMA) protocol are the three most commonly used. For the protocols described below, omnidirectional antenna is assumed. All stations use the same frequency band for transmission and hence stations are not allowed to transmit and receive at the same time.

1.3.1 Slotted ALOHA [ROBE 75]

The time axis is partitioned into fixed-size slots and it is assumed that the packet length fits within a slot. A station is allowed to transmit any time it wishes but only at the beginning of a slot. If the packet is collided, the station retransmits the packet after a random delay. This protocol is suitable for fixed-size packets but the channel utilization is low.

1.3.2 Carrier Sense Multiple Access [KLEI 75]

In this protocol a station is required to sense the channel before transmission. The station is allowed to transmit a packet only when no transmission is sensed. This protocol has several variants depending on the action taken when the channel is sensed busy. In the nonpersistent version the station would re-sense the channel after a random delay. In the persistent version the station would keep sensing the channel until it is idle and then transmit the packet. The CSMA protocol can offer superior performance than the ALOHA protocol. But its performance starts to degrade in a multihop network environment. This is mainly due to the hidden station problem. When a station A is transmitting a packet to its neighbouring station B, a station C, which is outside the transmission range of A but within range of B, cannot detect A's transmission by sensing the channel. Thus when C transmits, collision would occur at station B.

1.3.3 Busy Tone Multiple Access [TOBA 75]

The BTMA protocol can be used to solve the hidden station problem. In this protocol, a station broadcasts a busy tone in a separate channel whenever it is receiving a packet. A station is required to sense the busy tone channel before transmission and is not allowed to transmit during the presence of a busy tone. Collision, therefore, can be avoided. The additional busy tone bandwidth required is justified by the performance improvement.

1.4 SPATIAL REUSE

In a single-hop network using omnidirectional antennas, a packet transmission is received by all stations and therefore no simultaneous transmission is allowed. However, in a multihop network a packet transmission is received only by the neighbouring stations. Thus it is possible to reuse the same frequency and time in different parts of the network provided that the respective transmitting stations are far apart enough. The simultaneous use of the same channel bandwidth in different locations without interference is referred as spatial reuse. When a station uses a directional antenna for transmission, only a subset of neighbouring stations receives the packet. Thus spatial reuse is also possible in the single-hop environment. A survey of results on topological design and associated channel access protocols that attempt to optimize system performance by spatial reuse can be found in [KLEI 87].

The spatial reuse of the channel can gain more performance improvement only when suitable channel access protocol is used. Consider the CSMA protocol as an example. In the multihop environment, using carrier sensing protocols can only provide information about a station's local environment. Hearing the channel idle does not guarantee that the target receiver's environment is also idle and collision thus cannot be avoided. On the other hand sensing the channel busy and prohibiting packet transmission might reduce the benefit of spatial reuse. Consider the situation in Figure 1.1. Under the CSMA protocol, station C is not allowed to transmit when station A is transmitting a packet to station B. But station C actually could transmit a packet to station D without interfering station B's packet reception.

Another way of increasing the spatial reuse effect is by controlling the transmission power. A low transmission power causes less interference and leads to a higher degree of spatial reuse. But a packet would require a larger number of relay transmissions to reach the final destination. Thus the channel load is increased and the overall network performance therefore might not be improved. A high transmission power leads to a smaller number of transmission hops but causes greater interference. The Most Forward with Variable Radius (MVR) routing strategy [HOU 86] seems to be a compromise solution. In MVR once the repeater has been identified, a station would reduce the transmission power to exactly that needed. A high transmission power with adjusting capability is thus a desirable choice.

1.5 SPREAD SPECTRUM

Spread spectrum signaling opens up a new dimension for protocol design and performance trade-off. The signal capture, multiple access, anti-multipath and narrow-band interference rejection are the desirable characteristic that can be obtained via the use of spread spectrum [PURS 87]. The use of spreading codes permits a receiver to extract a particular signal from many overlapping ones. This kind of communications is called Code Division Multiple Access (CDMA). The use of a CDMA protocol allows simultaneous communication between many PRN stations through different codes.

It is difficult to design a receiver that can simultaneously monitor all the codes. Therefore there must be rules specifying which set of codes is to be monitored and which set of codes is to be used for transmission for each station. Four types of spreading code protocols can be identified: common code protocols, receiver-based protocols, transmitter-based protocols and hybrid protocols [SOUS 88].

With the common code approach, a single spreading code is used by all stations. As a result of the delay rejection capability of spread spectrum signals, it is possible to have simultaneous transmissions as long as there is a sufficiently large offset in code phases. In the receiver-based protocol a unique receiving code is assigned to each station. All packets are transmitted using the destination's code. In the receiving mode, a station constantly monitors its own code for detecting packets destinated to it. The drawback of this protocol is that collision still occurs when more than one packet is sent to the same destination. In the transmitter-based scheme a unique transmitting code is assigned to each station. Thus the transmissions of different stations would not interfere with each other. However this scheme creates a problem at the receiving end as a receiver cannot anticipate which stations is sending it packets. Hybrid protocols are formed from various combinations of the above three protocols in a fixed or dynamic fashion.

The characteristics of spread spectrum influence the choice of channel access protocols in PRNs and many new protocols have been designed. Two major factors of designing such protocols can be identified [PURS 87]. The first is to allow simultaneous transmissions as many as possible to achieve high network throughput. Thus channel sensing should be used to determine whether the intended receiver is busy instead of detecting other stations' transmissions within the local neighbourhood. Second, the way of channel sensing in narrow-band signaling cannot be directly applied in a spread spectrum network. For example, if the transmitter-based spreading code protocol is used, it is impossible for a station to monitor all the codes for detecting the channel status. The transmitter also cannot know which code should be monitored to determine whether the intended receiver is busy. Due to the above two factors, sometimes it is more efficient to use ALOHA type channel access protocols than to use channel sensing type protocols in Spread Spectrum Packet Radio Networks (SS/PRNs).

1.6 THESIS INTRODUCTION

We have introduced the network structure of a PRN and its components in this chapter. Three most commonly used channel access protocols, namely the Slotted ALOHA protocol, the CSMA protocol and the BTMA protocol, are described. We have also discussed the performance improvement gained by spatial reuse and the role of spread spectrum in PRNs. This brief overview of PRNs provides the background for the discussion in the following chapters.

In Chapter 2 we design a new protocol called the Simple Tone Sense (STS) protocol for multihop PRNs with multiple directional antennas stations. The protocol can minimize transmission interference by using a group of tones to identify the active neighbours. A variation of the STS protocol called the Variable Power Tone Sense (VPTS) protocol is also designed to further reduce interference. Algorithms for assigning tones and for determining the orientation and broadcasting angles of the directional antennas are designed. Design examples are given. Simulation result shows that the STS protocol gives better throughput-delay performance than the BTMA protocol, especially when the traffic is heavy. The VPTS protocol gives still better throughput-delay performance than the STS protocol.

In SS/PRNs not using common code approach, different spreading codes are required for different stations for transmitting packets. Therefore multihop SS/PRNs with a large number of stations would require a large number of codes and hence a large channel bandwidth. In Chapter 3 we design a code assignment algorithm which takes the spatial reuse into account and could reduce the number of codes required to about 20% of the total number of stations in the network. Further reducing the number of codes is found to cause little throughput degradation. The Coded Tone Sense protocol is designed for using these codes in multihop PRNs. Simulation result shows that in a 80 node network using only 5 spreading codes, the maximum network throughput is about 80% higher than that of the BTMA protocol.

Even with code reuse beyond the interference range, it is important to find an efficient algorithm for assigning as few codes to the SS/PRN stations as possible. In Chapter 4 we relate the code assignment problem to the graph coloring problem and propose a very efficient algorithm for assigning codes to the stations. A very tight lower bound on the number of codes needed is also derived.

In Chapter 5 we design a new scheduling algorithm for packet transmission in multihop PRNs. The design objective is to have a schedule with minimum cycle length, maximum network throughput and fair allocation of transmission capacities among all stations. For comparison with other scheduling algorithms in the literatures the following performance measures for scheduling algorithms are derived: (1) the cycle length, (2) the scheduling delay, (3) the minimum transmission capacity and (4) the normalized network capacity. The new algorithm is found to give schedules that have (1) the shortest cycle length, (2) the smallest scheduling delay, (3) the largest minimum transmission capacity and (4) the same highest normalized network capacity when compared to two of the best scheduling algorithms in the literature.

Broadcasting is very often used for updating distributed databases and routing tables in a communication network. In Chapter 6 we design the Staggered Multicast Protocol which is suitable for unicasting, broadcasting as well as multicasting in multihop SS/PRNs. The Common-Header/Transmitter-Based spreading code is used for data packets transmission and the receiver-based code is used for acknowledgement packets transmission. By staggering packet transmission the protocol can significantly reduce broadcasting delay. Special addressing method and packet format are also designed to achieve collision-free acknowledgement and multicast capability. Simulation result shows that the protocol provides better throughput-delay performance than the Common-Header/Transmitter-Based Slotted ALOHA protocol.

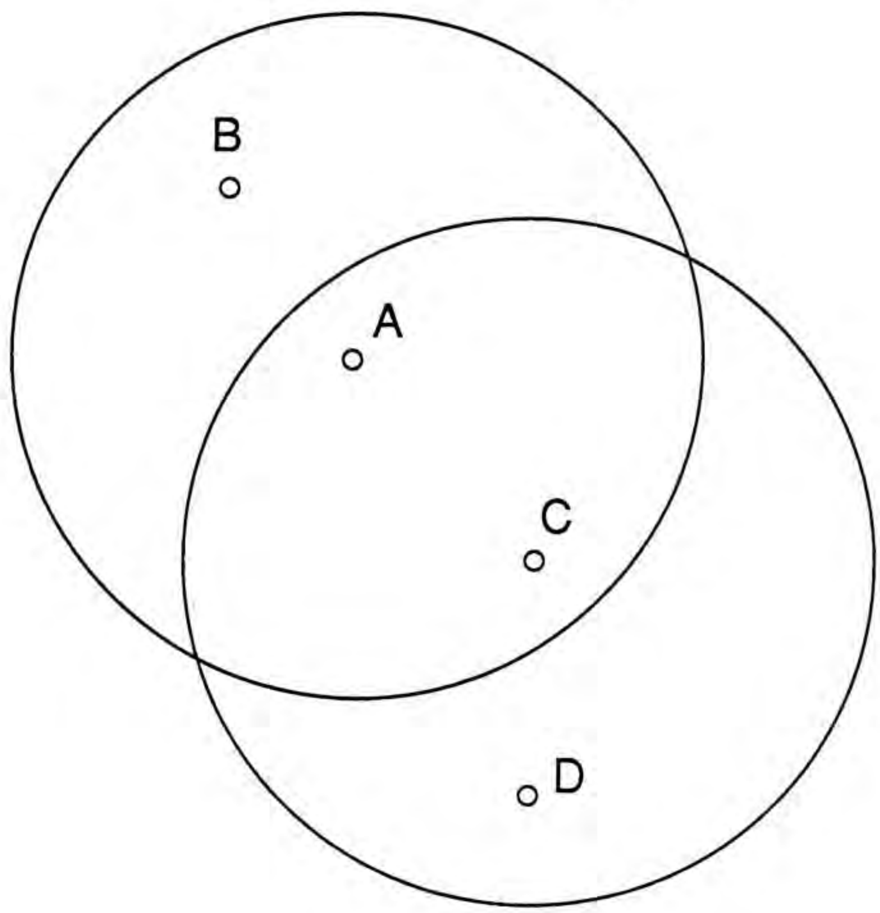


Fig.1.1 Spatial reuse effect

CHAPTER 2

DESIGN ALGORITHMS FOR NETWORKS WITH DIRECTIONAL ANTENNAS

2.1 INTRODUCTION

Omnidirectional antennas are usually used in the stations of the systems described in Chapter 1. In other research works, directional antennas are used to increase the network throughput by spatially reusing the packet radio channel. J.F. Chang and C.J. Chang examined the performance improvement of using directional antennas operated under the Slotted ALOHA protocol and the non-persistent CSMA protocol [CHAN 84]. They considered the system where each station was equipped with a single directional antenna only.

However, with a single directional antenna, there is a need to change the antenna directions for different receiving stations. To solve this problem, Hung and Yum proposed and analysed the performance of Slotted ALOHA with Multiple Directional Antenna (SA/MDA) [HUNG 86a] and found that the throughput and the expected progress are always higher than those using one omnidirectional antenna. When the number of stations in the transmission circle is fairly large, the gain could be as much as the number of directional antennas used.

Another example is the MTCD/MDA (Multi-Tone Multi-Access with Collision Detection using Multiple Directional Antennas) protocol [HUNG 86b]. In the MTCD/MDA protocol, when a station is busy in receiving packets each of its directional antennas broadcasts different busy tones. The orientation of the antennas must be the same for all stations and the number of antennas in each station must also be the same and be even.

With the use of directional antennas in multihop packet radio networks, a number of design issues arise. These include: 1) number of antennas per station, 2) orientation and broadcasting angle of each antenna, 3) how to minimize the interference between neighbouring stations, and 4) the design of efficient protocols to fully utilize the spatial frequency reuse advantage of directional antennas. In this chapter, we design a new protocol called the Simple Tone Sense (STS) protocol that can minimize transmission interference by using a group of tones to identify the neighbouring stations that are receiving packets. A variation of the STS protocol called the Variable Power Tone Sense (VPTS) protocol is also designed to further reduce interference by using minimum required transmission power to reach the intended neighbour. We first discuss the problems associated with the MTCD/MDA protocol in Section 2.2. In Section 2.3 and 2.4 we describe the multihop network environment and the new protocols to be used on it. Design algorithms for assigning tones and for determining the orientation and broadcasting angles of the directional antennas are presented in Section 2.5. A network design example is given in Section 2.6 and simulation results of the STS and VPTS protocols on four networks with various network size and density are given in Section 2.7.

2.2 PROBLEMS IN THE MTCD/MDA PROTOCOL

In the MTCD/MDA protocol, when a station detects a packet addressed to it, it broadcasts different busy tones to its neighbours in different directions. All stations must have the same number of antennas and the antenna orientation for all stations must be the same. These two restrictions cause the following problems in a multihop network:

(1) A large number of stations may be concentrated in certain transmission sectors, causing severe congestions in these sectors (Figure 2.1).

(2) Stations located near the quadrant boundaries are in the coverage of two antennas and suffered the interference from both antennas.

(3) Increasing the number of transmission sectors requires a corresponding increase in the number of tones.

(4) Since the number of antennas is the same for all stations, the station with more neighbours cannot have more antennas and the station with very few neighbours cannot have less. This problem is illustrated in Figure 2.2. Station A has twenty neighbouring stations while station B has only five.

A new protocol is proposed here that can avoid the above problems. The main idea is that each station is assigned a tone which is unique to its neighbours (We will describe how these tones are assigned in Section 2.5). When a station receives a packet, it broadcasts its tone immediately for a period of time so that its neighbours can identify its presence and avoid transmitting to its direction. With this arrangement, the orientation and broadcasting angles of the directional antennas are not required to be the same for all stations and the orientation of antennas can be chosen to make the neighbouring stations as far away from the boundary of the transmission sectors as possible. Moreover, the broadcasting angles of the antennas should be so chosen as to make the number of stations in each sector as evenly distributed as possible (Figure 2.3). Also, a station with more neighbours can be equipped with more directional antennas so that a particular antenna will not be too heavily loaded.

2.3 THE SIMPLE TONE SENSE (STS) PROTOCOL

2.3.1 System Descriptions

Let there be N stations in the packet radio network and let all station locations be fixed. Each station is assigned a tone and a station-number. A tone is just a sinusoidal wave at a certain frequency. The station-number is globally unique, but the tone frequency is unique only in each station's neighbourhood. The maximum transmission range R is assumed to be the same for all stations and the propagation delay across distance R be a.

Each station maintains five arrays to store the network status information. Consider a local station which has m directional antennas and n neighbouring stations with station-numbers 1, 2,..., n.

(a) Path Array $P = (p_1, p_2, ..., p_N)$ where p_j is the station-number of the neighbouring station that will lead to destination station j. For example, $p_{13} = 7$ means that a packet destined for station 13 is to be sent via station 7.

(b) Antenna Array $A = (a_1, a_2, ..., a_n)$ where a_k denotes the directional antenna for transmitting packets to station k. For example, $a_4 = 3$ means that station 4 is in the coverage of the directional antenna 3 of the local station.

(c) Neighbour Status Array $S = (s_1, s_2, ..., s_n)$ shows the Busy/Idle status of the neighbours of the local station. Thus, $s_6 = 0$ and $s_{10} = 1$ indicate that stations 6 and 10 are idle and busy respectively.

(d) Antenna Status Array $B = (b_1, b_2, ..., b_m)$ shows the Busy/Idle status of the m directional antennas of the local station with a 0 indicating idle and a 1 indicating busy.

(e) Sector Status Array $C = (c_1, c_2, ..., c_m)$ shows the Busy/Idle status of the m transmission sectors of the local station. The value of c_i is determined from the busy/idle status of the stations (i.e. the array S) covered by directional antenna i. For example, $c_3 = 0$ means that all stations in sector 3 (covered by directional antenna 3) are idle while $c_3 = 1$ means that one or more stations in sector 3 are currently receiving packets.

Tone signaling serves both as an explicit destination to source acknowledgment and as a busy signal to alert other stations that a successful packet reception is ongoing. As soon as a station detects a packet destined to it, the station broadcasts its assigned tone for T_1 seconds to all its neighbours. As an example of the signaling and acknowledging process, consider the case in Figure 2.4. Here, A (or station A) is transmitting a packet to B , C knows that D's sector is idling and transmits a packet to D. But D cannot receive the packet because it is being interfered by A's transmission. Therefore C stops its transmission as soon as the expected tone from D is not received in a time out period of T_o seconds. Here T_o must be at least the round-trip propagation delay $2a$ plus the header detection time.

All stations when detecting a tone of duration T_1 will change the status of the station corresponding to the received tone to busy. After a station has received a packet correctly, it acknowledges the source station by broadcasting its assigned tone for T_2 seconds. All stations when detecting a tone for T_2 seconds will change the status of the corresponding station to idle. We will discuss how to determine the tone durations T_1 and T_2 in Section 2.5.

2.3.2 Transmission Protocol

(1) Look up the Path Array to decide which neighbour the packet is to be forwarded.

(2) Look up the Antenna Array to decide the appropriate directional antenna to be used and denote it as antenna k.

(3) If $b_k = 1$ or antenna k is being used at that moment, wait until it becomes idle.

(4) If $c_k = 1$ or some other station in sector k is receiving a packet, recheck after a random delay. If $c_k = 0$, transmit the packet immediately.

(5) If the expected tone of duration T_1 is detected before the time out period, continue the transmission. Otherwise, stop transmission immediately, wait for a random delay and go to (4).

(6) After the whole packet is transmitted, if the expected tone of duration T_2 is detected before the time out period, the packet is assumed to be correctly received. Otherwise, retransmit the packet immediately and go to (5).

2.3.3 Reception Protocol

(1) When the header of an incoming packet is detected without error, broadcast the assigned tone to all neighbours for T_1 seconds.

(2) After receiving the whole packet and no error is detected, broadcast the assigned tone for T_2 seconds.

2.4 THE VARIABLE POWER TONE SENSE (VPTS) PROTOCOL

In order to increase the spatial-reuse advantage, a station uses only the minimum required power to send the packet to the destination station. However, the assigned tones are still broadcasted to cover a range of R.

To illustrate the advantage of using variable power, consider the case in Figure 2.5. Here C transmits a packet to B with the minimum required power. B then broadcasts its tone for T_1 seconds. A detects the tone from B and will refrain from sending packets to stations B, D, E and F if the Simple Tone Sense protocol were used. But in reality, A could transmit a packet to D using the minimum required power without interfering B's packet reception. Now suppose E transmits a packet to B. A detects B's tone and realizes that B is far enough away from D. Therefore A might transmit a packet to D. But here, since D is in the transmission range of E, a collision will occur. However A's transmission cannot reach B, therefore B's packet reception is not affected. Thus the VPTS protocol can in some cases improve the network throughput but in no case will it degrade the throughput.

For each station, two additional arrays besides the five required by the STS protocol are maintained:

(a) Distance Array D = (d_1, d_2,..., d_n) where d_j is the distance between the local station and its j-th neighbour.

(b) Range Array $E = (r_1, r_2, \ldots, r_m)$ where r_k is the current allowable transmission range in sector k.

The Range Array is updated as follows:

(1) Initially all r_k's are set to R_o which is an arbitrary number larger than R.

(2) When the local station detects a tone of duration T_1 from neighbour j located in sector k, it replaces r_k by d_j if $d_j < r_k$.

(3) When the local station detects a tone of duration T_2 from its neighbour located in sector k, it determines from arrays S and D the nearest neighbour (if any) in sector k which is still receiving a packet and denote it as neighbour i. Set r_k to d_i.

(4) If no station in sector k is currently receiving a packet, r_k is set to R_o.

The transmission protocol of VPTS is identical to STS except that step (4) is replaced by:

(4) If $r_k > d_j$, where j is the intended receiving neighbour, transmit the packet immediately. Otherwise, recheck after a random delay.

This protocol assumes that stations can make continuous adjustment of transmission power, we call it the continuous range VPTS protocol. A discrete range version of the VPTS protocol can similarly be defined.

2.5 NETWORK DESIGN ALGORITHMS

2.5.1 Tone Assignment Algorithm

Each station is assigned a tone for identifying itself from the neighbouring stations. Since the identifying tones are local in nature, beyond a certain range, which we call it the local-range for convenience, they can be reused. A tone-group is just a group of stations that use the same tone for identifying themselves in their neighbourhood. The size of the local-range depends on the transmission range R and the distribution of neighbouring stations. The local-range of a particular station is formed by the perimeters of the transmission ranges of the station's neighbours (Figure 2.6).

The algorithm for assigning tones to stations is as follows:

(1) j := 1.

(2) Select an unassigned station and denote it as S_o.

(3) Assign tone j to S_o.

(4) S := S_o.

(5) Mark all the stations in the local-range of station S.

(6) If all unassigned stations are marked, go to (9).

(7) Assign tone j to one of the unmarked stations S'.

(8) S := S' ; go to (5).

(9) If all stations are assigned, stop. Otherwise unmark all unassigned stations.

(10) j := j + 1 ; go to (2).

2.5.2 Directional Antenna Assignment Algorithm

Each station may have a different number of directional antennas and each antenna can have a different broadcasting angle. The orientations and the broadcasting angles of the directional antennas are so chosen as to make the number of stations in each transmission sector as even as possible. Increasing the number of the directional antennas in the stations can improve the network throughput without increasing the number of tones required.

To avoid interference, let φ be the minimum angular separation required between a neighbouring station and the boundaries of the transmission sector (Figure 2.7(a)). Also let the nominal number of stations covered by an antenna be n_o. The neighbouring stations are assigned to the appropriate antennas according to the following three criteria:

(1) If two adjacent neighbouring stations have angular separation smaller than 2φ, both are assigned to the same antenna (Figure 2.7(b)). This will ensure that all stations are at least φ degrees away from the antenna transmission boundaries.

(2) The number of stations in each sector should be as even as possible and should normally not exceed n_o except when the neighbouring stations are very close together.

(3) A minimum number of directional antennas is preferred.

The algorithm for determining the beam width and the orientation of the directional antennas of any station, say station A, is as follows :

(1) Rank the neighbours of station A according to their bearings.

(2) Calculate the angles θ_1, θ_2,... between all adjacent neighbours of station A (Figure 2.7(b)).

(3) Partition the neighbouring stations into groups such that the angular separation between a station and its "closest" (i.e. with the smallest angular separation) group member is less than 2φ.

(4) Let P be the initial grouping after partitioning. Let the number of groups in P be p.

(5) Let G be the desirable grouping of the neighbours and g be the number of groups in G.

(6) G := P and g := p.

(7) k := 1.

(8) k' := k.

(9) If the total number of stations in group k' and its adjacent group k'+1 (mod k) is less than or equal to n_o, combine the two groups and call it group k'+1.

(10) If k' $\neq$ k - 1 (mod k), k' := k'+1 and go to (9).

(11) Denote the new grouping of neighbours as H and the number of groups in H as h. Let the number of stations in these groups be $n_1, n_2, \ldots, n_h$ and let Var(H) be the variance of the n_i's, or $Var(H) = \sum_{i=1}^{h} n_i^2 - \left(\sum_{i=1}^{h} n_i \right)^2$.

(12) If (h < g) OR (h = g and Var(H) < Var(G)) then G := H and g := h.

(13) If k < p, k := k+1 and go to (8).

(14) If k = p, the assignment is complete.

As an example, consider Figure 2.8 which shows station 2 and its neighbours. The neighbours are ordered as 8, 11, 23, 16, 5, 91, 64, 24, 50, 36, 74, 29, 47, 18, 32 according to their bearings from station 2. The angular separation θ's for all neighbour pairs are computed. Here, θ_1 and θ_2 are less than 2φ, but θ_3 and θ_{15} are greater than 2φ. Hence stations 8, 11 and 23 are assigned to the same group (group 1). The other groups can be obtained similarly and the grouping are shown as follows:

Group number	Station number	Number of stations in the group
1	8, 11, 23	3
2	16	1
3	5, 91	2
4	64	1
5	24, 50, 36, 74, 29	5
6	47, 18	2
7	32	1

The number of initial groups p = 7 and the number of stations in groups 1 to 7 are 3, 1, 2, 1, 5, 2 and 1 respectively. Suppose the nominal number of stations in an antenna zone n_o is four. Starting from group 1, we obtain the total number of stations in groups 1 and 2 is four, which is equal to n_o. So we combine the two groups and denote it as group 2. After that we check the total number of stations in groups 2 and 3 and find that it is 4+2 = 6 which is greater than n_o. So, groups 2 and 3 are not combined. The entire grouping process starting from group 1 is shown as follows:

Step	Grouping						
-	3	1	2	1	5	2	1
1		4	2	1	5	2	1
2		4	2	1	5	2	1
3		4		3	5	2	1
4		4		3	5	2	1
5		4		3	5	2	1
6		4		3	5		3

The result of the grouping is H = [4, 3, 5, 3] and the number of groups h is 4. The same steps are performed starting from different initial groups. The result of all possible grouping is shown as follows:

Starting group	Resultant sequence						
1	-	4	-	3	5	-	3
2	3	-	-	4	5	-	3
3	-	4	-	3	5	-	3
4	-	4	-	3	5	-	3
5	-	4	-	3	5	-	3
6	-	4	-	3	5	-	3
7	4	-	-	4	5	2	-

Since the minimum number of groups g is 4, the grouping G that we choose for station 2 is [4, 3, 5, 3] (which has a smaller variance compared with the grouping [4, 4, 5, 2]). The final orientation and broadcasting angles of the antennas are shown in Figure 2.9.

2.5.3 Routing Strategy

In a distributed multihop PRN, the choice of a routing algorithm is essential. We choose, for our design, the minimum hop routing rule for simplicity. When there are multiple minimum hop paths between two stations, one of them is chosen arbitrary. Let $p_{i,j}$ be the next station on the routing path from station i to station j. The routing path can be uniquely determined by the NxN path matrix $P = [p_{i,j}]$.

2.5.4 Tone Detection Time and Packet Length

When two or more stations transmit packets simultaneously to the same destination, the destination will experience a collision. Hence a busy tone should be broadcast only after the destination has not sensed a collision for a seconds. As collision is detected by checking the CRC code in the packet header, packet header transmission time T_h must be longer than a, or

$$T_h > a. \tag{2.1}$$

To make the protocol more efficient, packet transmission time T_p should be longer than the busy tone duration T_1 plus the round-trip propagation delay $2a$ or

$$T_p > T_1 + 2a \tag{2.2}$$

because otherwise gaps will always appear between successive packet transmissions and the acknowledgement tone will interfere with the busy tone.

To determine T_1 and T_2, consider the following cases. Let station i transmits a packet to station j at t_o. Then station j can receive the packet on or before t_o+a and broadcast the assigned tone immediately. This tone can reach all of station j's neighbours before t_o+2a. Hence, if a neighbouring station transmits a packet towards station j's direction in $[t_o, t_o+2a]$, a collision will occur.

Let us say in the worst case, one of station j's neighbouring stations transmits a packet at $t_o + 2a - \varepsilon$ (ε is arbitrarily small) or just before the tone arrives. That packet will reach station j at t_o+3a and cause a collision. To make sure that this collision can be detected, the tone duration T_1 must be at least $(t_o+3a) - (t_o+a) = 2a$ seconds (Figure 2.10) or

$$T_1 > 2a. \tag{2.3}$$

To distinguish T_2 from T_1, T_2 needs only be longer than T_1 or

$$T_2 > T_1. \tag{2.4}$$

As an example, let the transmission range R be 3 Km and the data rate be 500 Kbps. Then $a = 0.01$ ms and the header length need only be longer than 5 bits. If we choose T_1 to be 2 ms which satisfies constraint (2.3), then the minimum packet size from constraint (2.2) is 2.02 ms, or 1010 bits.

Note that a tone is just a pure sine wave at a certain frequency and theoretically occupies zero bandwidth. But the turning-on and turning-off of a busy tone make the tone signal look like an On-Off Keying signal. The shortest duration of a tone in our case is T_1. So the tone bandwidth is about $2/T_1$ Hz. To minimize the bandwidth occupied by the tone, T_1 therefore should be as large as possible. For a data rate of 500 Kbps, the data bandwidth is 1 MHz. If T_1 is chosen as 2 ms, the tone bandwidth is only 1 KHz. For a system using 20 tones, the tones occupy a bandwidth of 20 KHz. The total bandwidth needed is therefore only 1.02 MHz.

2.6 NETWORK DESIGN EXAMPLE

Consider a PRN with 20 randomly located stations on a 20km x 20 km square area. Let the transmission range R be 8 km. A particular sampling gives the following station locations:

Station-number	1	2	3	4	5	6	7	8	9	10
X coordinate	8.4	16.1	13.9	3.2	1.6	3.2	18.6	14.5	14.3	17.8
Y coordinate	3.3	15.4	18.8	17.8	13.5	3.1	0.2	15.1	14.2	0.4

Station-number	11	12	13	14	15	16	17	18	19	20
X coordinate	1.0	16.5	13.1	16.3	7.8	5.7	4.6	15.4	5.5	18.8
Y coordinate	13.1	8.0	17.1	9.2	9.0	18.0	13.9	10.9	19.5	6.8

The tone assignment according to the algorithm is:

Station-number	1	2	3	4	5	6	7	8	9	10
Tone-number	1	1	2	1	2	3	1	3	4	2

Station-number	11	12	13	14	15	16	17	18	19	20
Tone-number	4	5	6	7	8	5	7	9	10	10

Next, the nominal number of stations covered by a directional antenna and the angular interference margin φ are set to be three and 0.2 radian respectively. Following the antenna assignment algorithm, the antenna coverage matrix $A = [a_{ij}]$ is determined. Here, $a_{ij} = k$ means that station j is in the coverage of the directional antenna k of station i and $a_{ij} = 0$ means station j is not a neighbour of station i. The matrix A therefore also contains the full connectivity information of the network.

$$
A = \begin{matrix}
0&0&0&0&0&1&0&0&0&0&0&0&0&0&1&0&0&0&0&0 \\
0&0&1&0&0&0&0&2&2&0&0&3&1&3&0&0&0&3&0&0 \\
0&1&0&0&0&0&0&2&2&0&0&0&1&0&0&0&0&0&0&0 \\
0&0&0&0&2&0&0&0&0&0&2&0&0&0&0&1&2&0&1&0 \\
0&0&0&2&0&0&0&0&0&0&1&0&0&0&1&2&1&0&2&0 \\
1&0&0&0&0&0&0&0&0&0&0&0&0&0&1&0&0&0&0&0 \\
0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&0&0&1 \\
0&1&1&0&0&0&0&0&2&0&0&3&2&3&0&0&0&3&0&0 \\
0&1&2&0&0&0&0&2&0&0&0&3&2&3&0&0&0&3&0&0 \\
0&0&0&0&0&0&1&0&0&0&0&1&0&0&0&0&0&0&0&1 \\
0&0&0&1&1&0&0&0&0&0&0&0&0&0&2&1&1&0&1&0 \\
0&1&0&0&0&0&0&1&1&2&0&0&0&1&0&0&0&1&0&2 \\
0&1&1&0&0&0&0&3&3&0&0&0&0&0&0&2&0&3&2&0 \\
0&1&0&0&0&0&0&1&1&0&0&2&0&0&0&0&0&1&0&2 \\
1&0&0&0&2&1&0&0&0&0&2&0&0&0&0&0&1&1&0&0 \\
0&0&0&1&2&0&0&0&0&0&2&0&1&0&0&0&2&0&1&0 \\
0&0&0&1&2&0&0&0&0&0&2&0&0&0&2&1&0&0&1&0 \\
0&1&0&0&0&0&0&1&1&0&0&3&1&3&2&0&0&0&0&3 \\
0&0&0&1&1&0&0&0&0&0&1&0&2&0&0&2&2&0&0&0 \\
0&0&0&0&0&0&2&0&0&2&0&1&0&1&0&0&0&1&0&0
\end{matrix}
$$

Finally, the path matrix P is constructed according to the routing rule as

$$P = \begin{matrix}
0 & 15 & 15 & 15 & 15 & 6 & 15 & 15 & 15 & 15 & 15 & 15 & 15 & 15 & 15 & 15 & 15 & 15 & 15 & 15 \\
18 & 0 & 3 & 13 & 18 & 18 & 12 & 8 & 9 & 12 & 18 & 12 & 13 & 14 & 18 & 13 & 13 & 18 & 13 & 14 \\
9 & 2 & 0 & 13 & 13 & 13 & 8 & 8 & 9 & 8 & 13 & 8 & 13 & 8 & 13 & 13 & 13 & 8 & 13 & 2 \\
17 & 16 & 19 & 0 & 5 & 11 & 19 & 19 & 16 & 11 & 11 & 16 & 19 & 11 & 11 & 16 & 17 & 19 & 19 & 11 \\
15 & 19 & 16 & 4 & 0 & 15 & 15 & 15 & 16 & 15 & 11 & 15 & 16 & 15 & 15 & 16 & 17 & 15 & 19 & 15 \\
1 & 15 & 15 & 15 & 15 & 0 & 15 & 15 & 15 & 15 & 15 & 15 & 15 & 15 & 15 & 15 & 15 & 15 & 15 & 15 \\
20 & 10 & 20 & 20 & 20 & 20 & 0 & 20 & 20 & 10 & 20 & 10 & 20 & 20 & 20 & 20 & 20 & 20 & 20 & 20 \\
18 & 2 & 3 & 13 & 13 & 18 & 18 & 0 & 9 & 12 & 18 & 12 & 13 & 14 & 18 & 13 & 18 & 18 & 13 & 12 \\
18 & 2 & 3 & 13 & 13 & 18 & 18 & 8 & 0 & 12 & 18 & 12 & 13 & 14 & 18 & 13 & 18 & 18 & 13 & 12 \\
12 & 12 & 12 & 20 & 12 & 12 & 7 & 12 & 12 & 0 & 12 & 12 & 20 & 20 & 12 & 20 & 20 & 20 & 20 & 20 \\
15 & 19 & 19 & 4 & 5 & 15 & 15 & 16 & 19 & 15 & 0 & 15 & 19 & 15 & 15 & 16 & 17 & 15 & 19 & 15 \\
18 & 2 & 2 & 18 & 18 & 18 & 20 & 8 & 9 & 10 & 18 & 0 & 2 & 14 & 18 & 8 & 18 & 18 & 18 & 20 \\
18 & 2 & 3 & 16 & 19 & 18 & 18 & 8 & 9 & 9 & 16 & 18 & 0 & 8 & 18 & 16 & 19 & 18 & 19 & 18 \\
18 & 2 & 9 & 8 & 18 & 18 & 20 & 8 & 9 & 12 & 18 & 12 & 8 & 0 & 18 & 6 & 18 & 18 & 2 & 20 \\
1 & 18 & 18 & 17 & 5 & 6 & 18 & 18 & 18 & 18 & 11 & 18 & 18 & 18 & 0 & 11 & 17 & 18 & 11 & 18 \\
17 & 13 & 13 & 4 & 5 & 17 & 13 & 13 & 13 & 13 & 11 & 13 & 13 & 13 & 17 & 0 & 17 & 13 & 19 & 13 \\
15 & 19 & 16 & 4 & 5 & 15 & 15 & 15 & 16 & 15 & 11 & 15 & 16 & 15 & 15 & 16 & 0 & 15 & 19 & 15 \\
15 & 2 & 8 & 13 & 15 & 15 & 20 & 8 & 9 & 20 & 15 & 12 & 13 & 14 & 15 & 13 & 15 & 0 & 13 & 20 \\
5 & 13 & 13 & 4 & 5 & 11 & 13 & 13 & 13 & 13 & 11 & 13 & 13 & 13 & 11 & 16 & 17 & 13 & 0 & 13 \\
18 & 18 & 18 & 18 & 18 & 18 & 7 & 12 & 18 & 10 & 18 & 12 & 18 & 14 & 18 & 18 & 18 & 18 & 18 & 0
\end{matrix}$$

2.7 SIMULATION RESULTS

Four network samples are generated on which the performance of various protocols are compared. The stations in the network are randomly located within a 20 km x 20 km square region. The transmission range is 4 km. The nominal number of stations covered by a directional antenna and the angular interference margin are 5 and 0.2 radian respectively. The packet generation rates are the same for all stations and the packet destinations are equally probable for all stations, excluding the source station. Let the packets be of fixed length and let the arrivals to each station be a Poisson process. The characteristics of the networks generated are summarized as follows:

Network parameters \ Cases	1	2	3	4
No. of stations	80	80	40	40
ave. no. of neighbours per station	8.31	8.38	4.05	5.45
max. no. of neighbours per station	15	12	7	11
ave. no. of antennas per station	2.10	1.89	1.23	1.55
max. no. of antennas per station	4	3	2	3
No. of tones required	21	15	9	14

The protocols compared include Slotted ALOHA with single omnidirectional antenna (SA), Slotted ALOHA with multiple directional antennas (SA/MDA), Busy Tone Multiple Access with omnidirectional antenna (BTMA), Simple Tone Sense (STS) and Variable Power Tone Sense (VPTS) protocols. Minimum hop routing rule is used. The normalized network throughput, or the average number of packets reaching destinations per packet transmission time is measured in the simulation. This throughput measure is different from the *one-hop throughput* usually given in some studies because most packets have to travel two or more hops before reaching their destinations. We assume the total bandwidth occupied by the busy tones is 2% of the total bandwidth for STS and VPTS protocols. The throughput η shown for these two protocols is the effective network throughput, which is the network throughput multiply by (1 - 0.02).

The average end-to-end delay as a function of network throughput for cases 1 and 3 are plotted in Figures 2.11 and 2.12 respectively. The throughput-delay characteristics of cases 2 and 4 are similar and so are not shown. The maximum network throughput attained for the four station distributions (or the four cases) are obtained as follows:

Protocols \ Cases	1	2	3	4
SA	0.50	-	0.35	-
SA/MDA	0.71	0.69	0.40	0.39
BTMA	1.09	1.08	0.78	0.79
STS	1.42	1.57	0.79	0.80
VPTS	1.57	1.74	0.98	0.88

The maximum network throughput attained by STS has 30% to 45% improvement over BTMA on the 80 node networks. This improvement is primarily due to the use of additional directional antennas. But on the 40 node networks there is no significant improvement. Therefore, it seems that when the stations are not densely located, adding additional directional antennas does not significantly improve the network throughput. But the delay for STS and VPTS is always smaller than BTMA. When compare VPTS with STS, there is a throughput improvement around 10% on the 80 node networks. On the 40 node networks, the improvement in network throughput is 24% for case 3 and 10% for case 4. VPTS always has a smaller delay than STS.

2.8 CHAPTER SUMMARY

Most Packet Radio Network protocols are designed for stations with omnidirectional antennas. This is due mainly to the simplicity requirement of the system and the protocol. Using directional antennas however has the advantage of greater spatial reuse of the radio channel and leads to a higher throughput of the network. This chapter is an attempt to give a design methodology as well as two efficient transceiving protocols for multihop Packet Radio Networks with multiple directional antennas stations. We have designed the Simple Tone Sense (STS) protocol which can minimize transmission interference by using a group of tones to identify the active neighbouring stations. A variation of the STS protocol namely the Variable Power Tone Sense (VPTS) protocol has also been designed to further reduce interference. Algorithms for assigning tones and for determining the orientation and broadcasting angles of the directional antennas have also been designed. Simulation result shows that the STS protocol performs better than the BTMA protocol, especially when the traffic is heavy. But the VPTS protocol gives still better throughput-delay performance than the STS protocol.

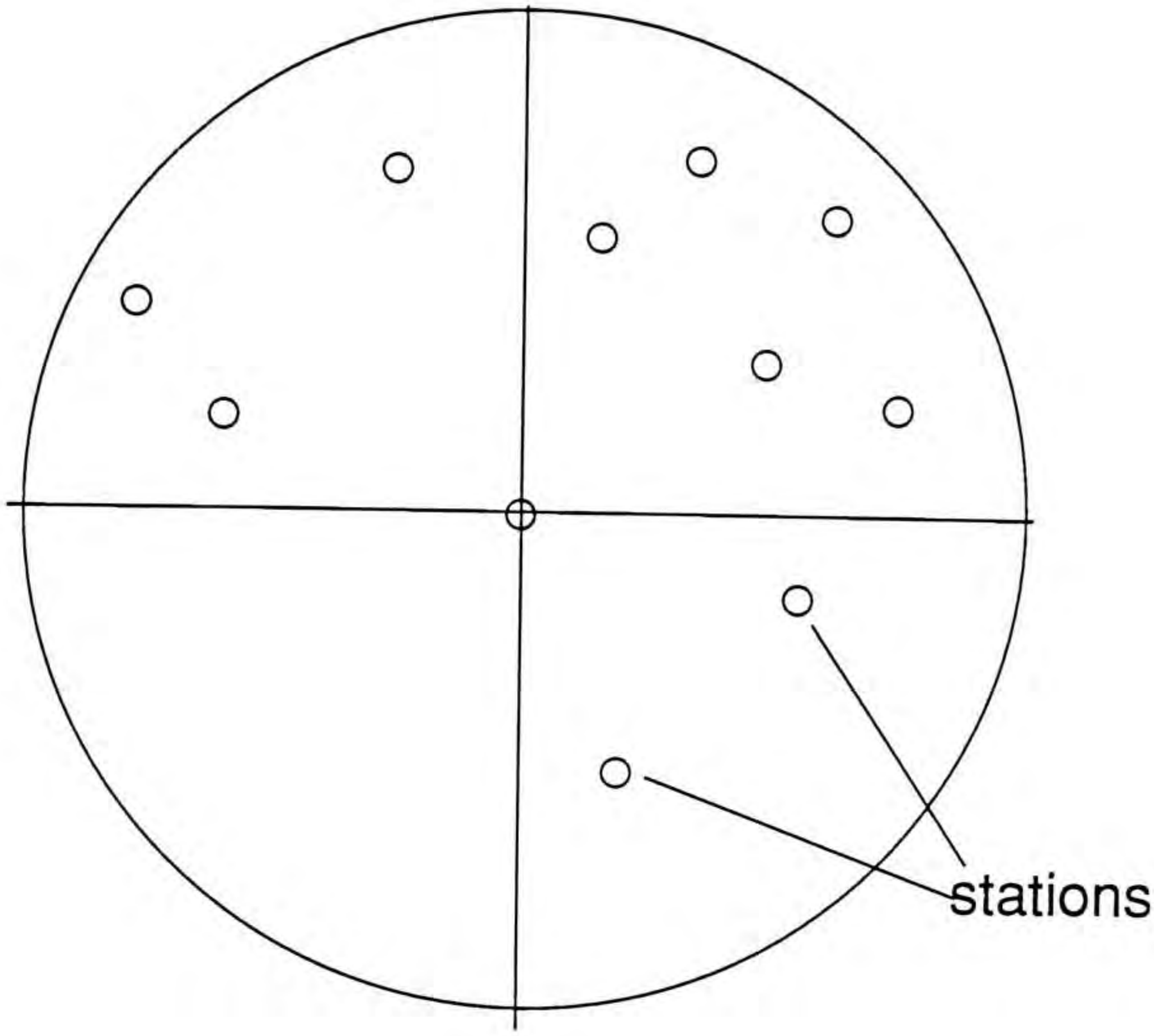


Fig.2.1 Limitations of the MTCD/MDA protocol

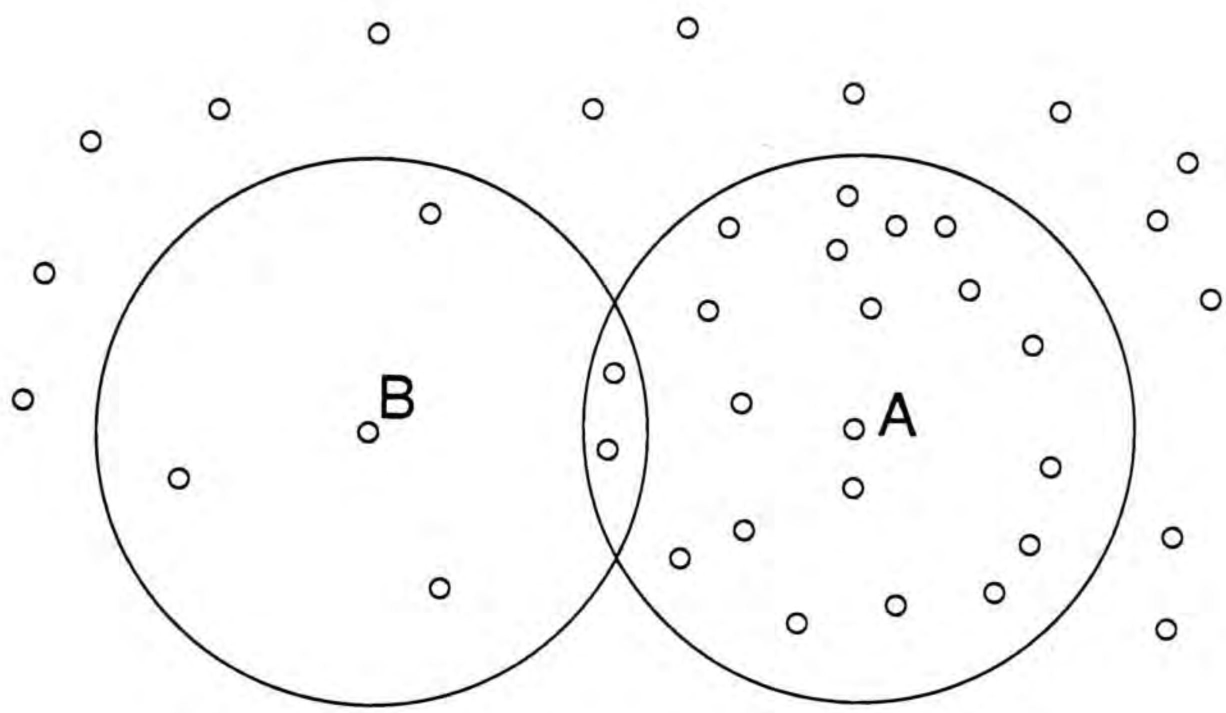


Fig.2.2 A distributed multihop network

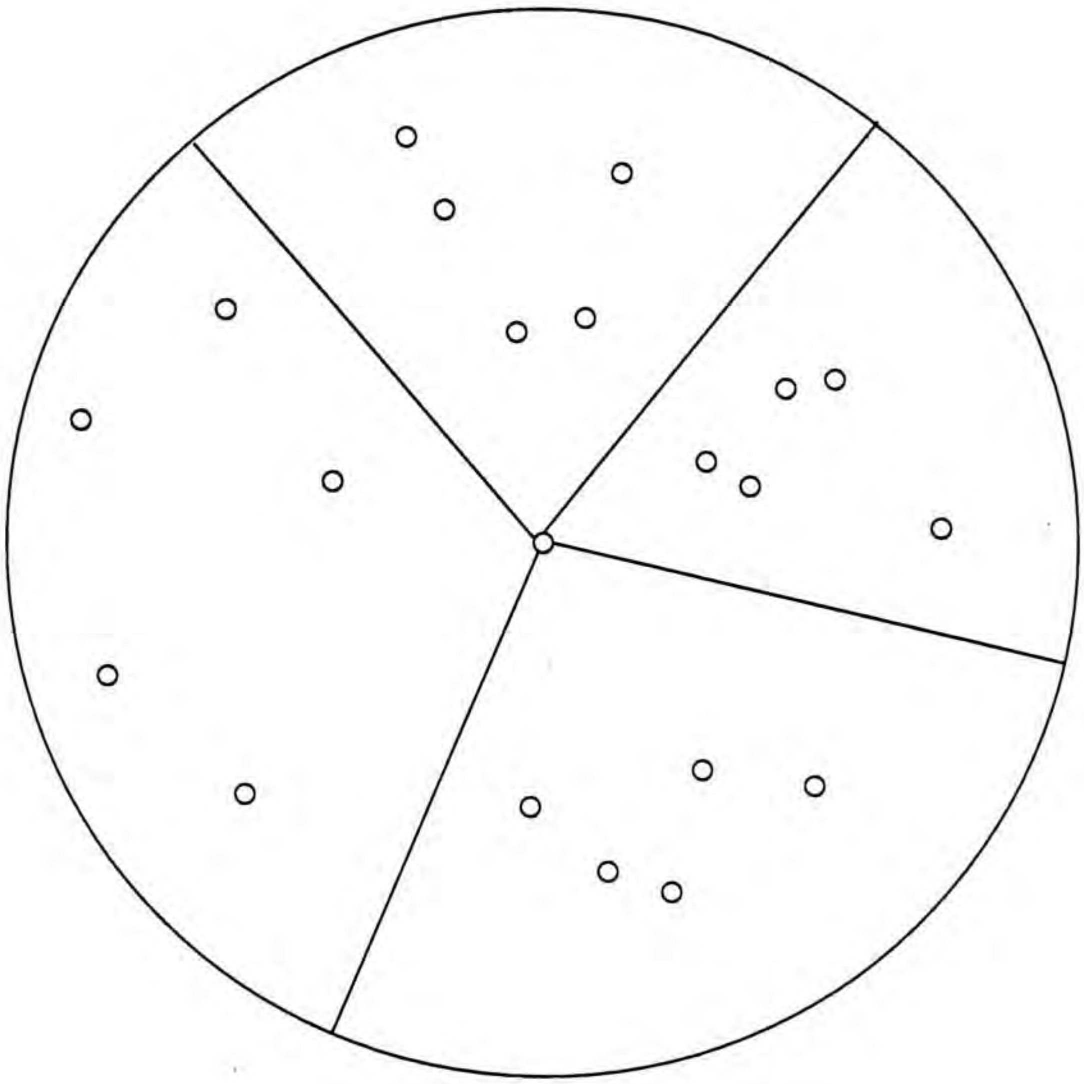


Fig.2.3 Even distribution of stations among sectors

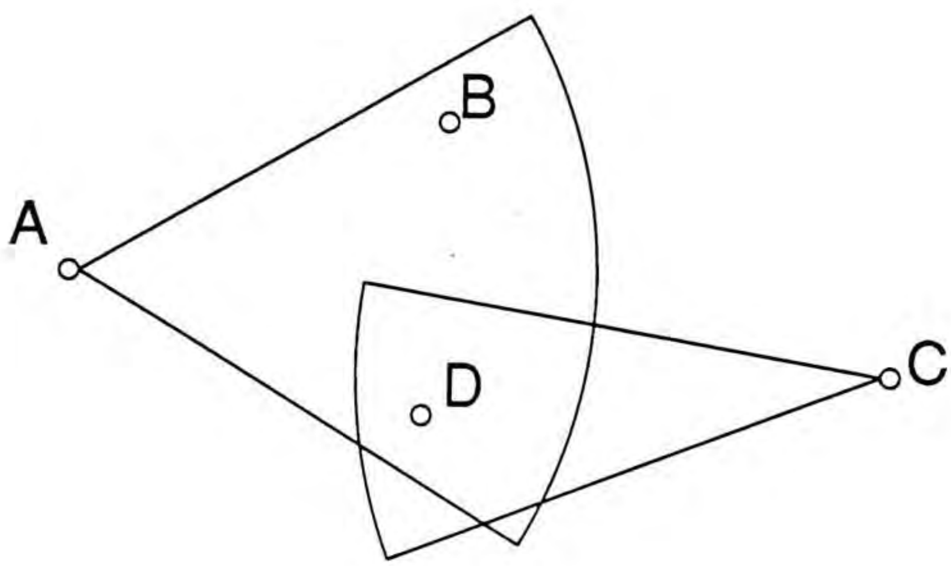


Fig.2.4 An example of signaling and acknowleging in STS

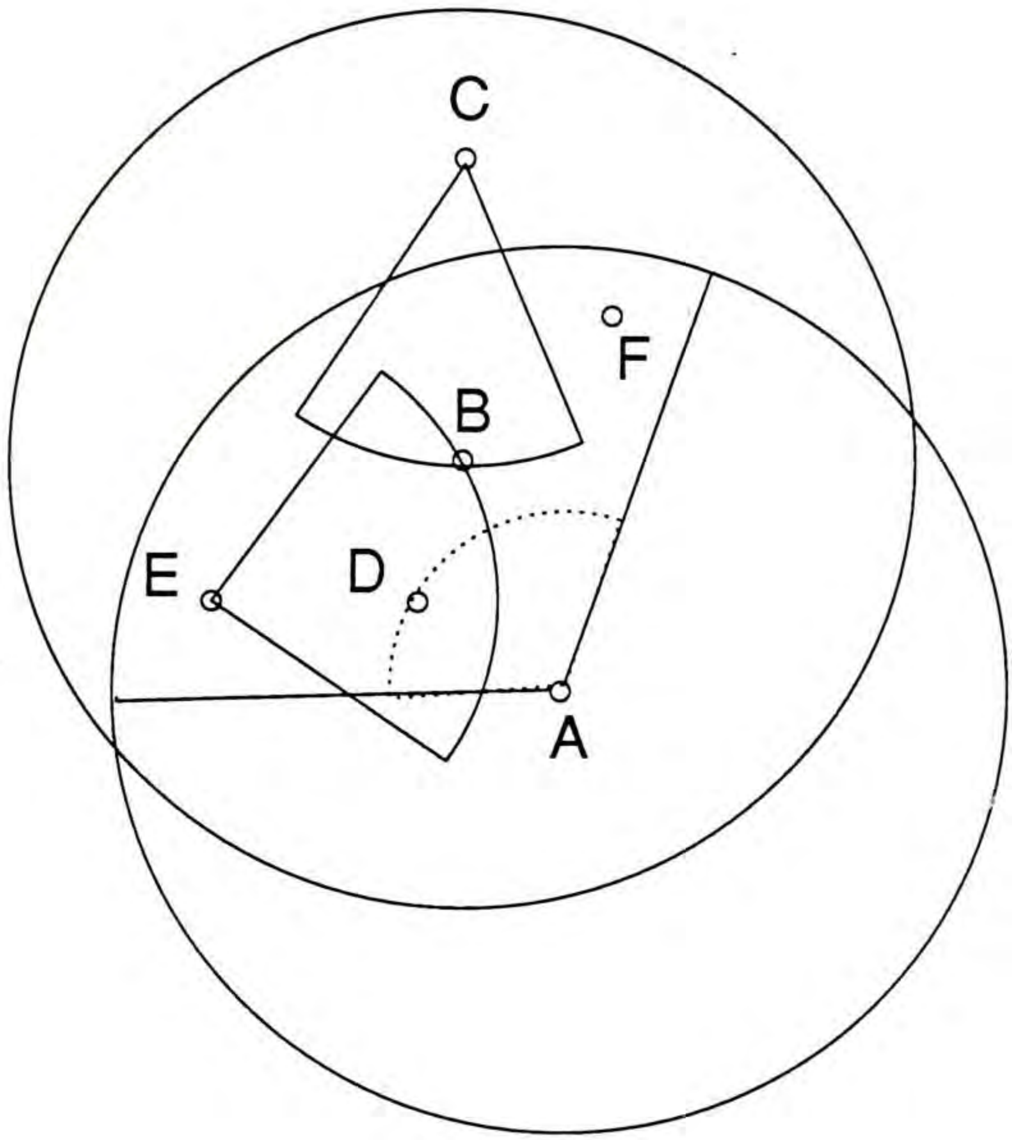


Fig.2.5 Variation of transmission power

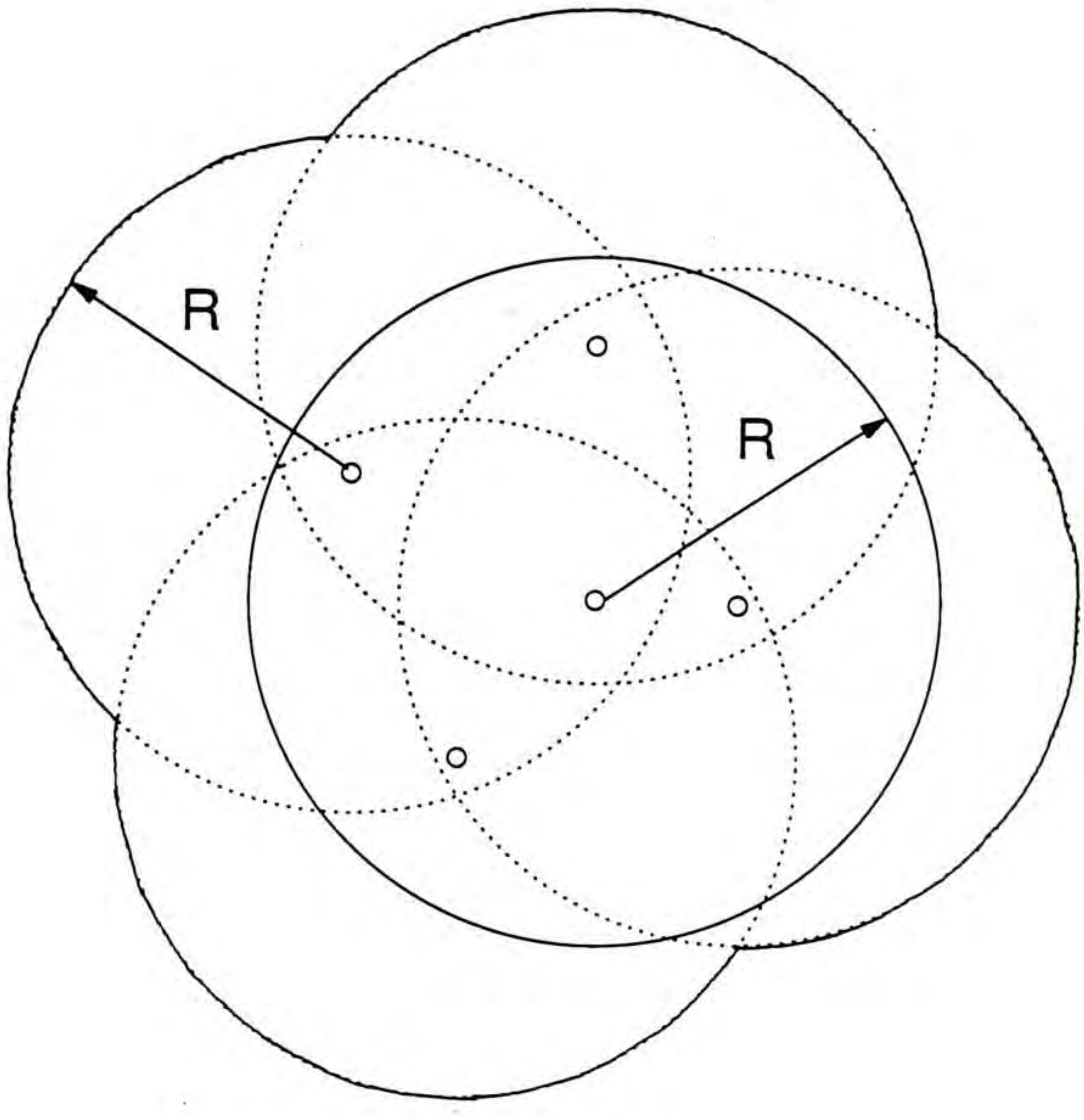
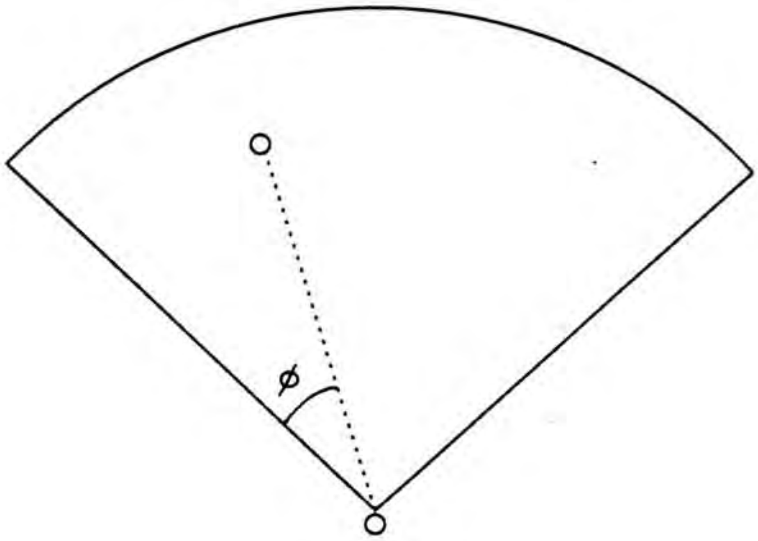
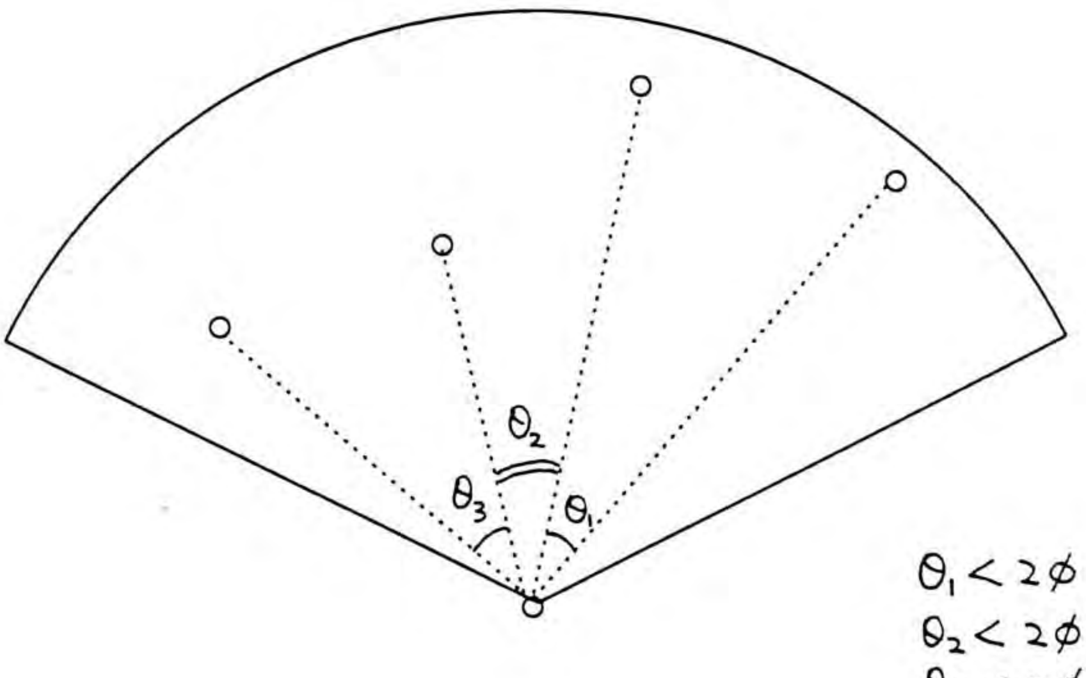


Fig.2.6 The local range



(a)



(b)

Fig.2.7 Antenna assignment

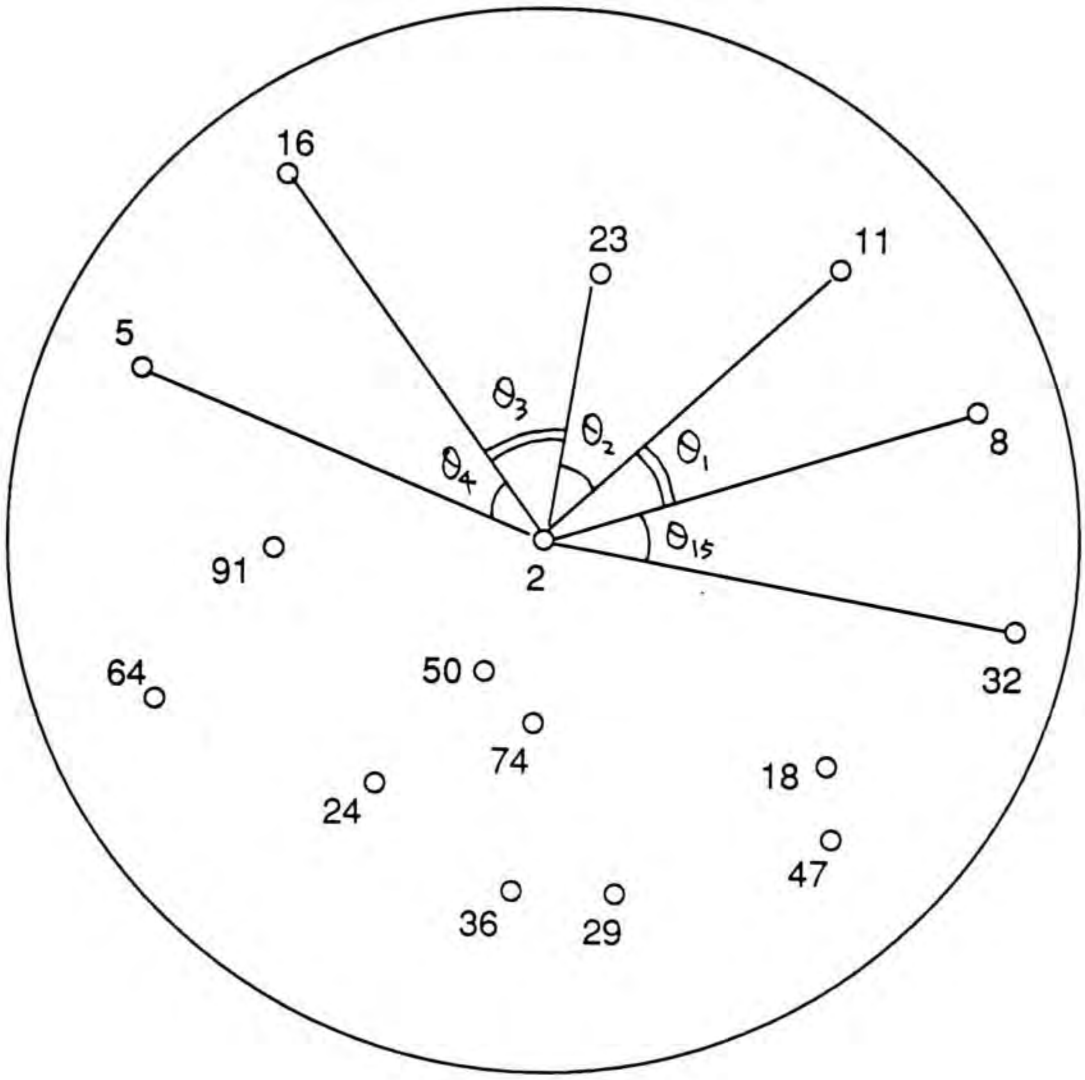


Fig.2.8 Station 2 and its neighbours

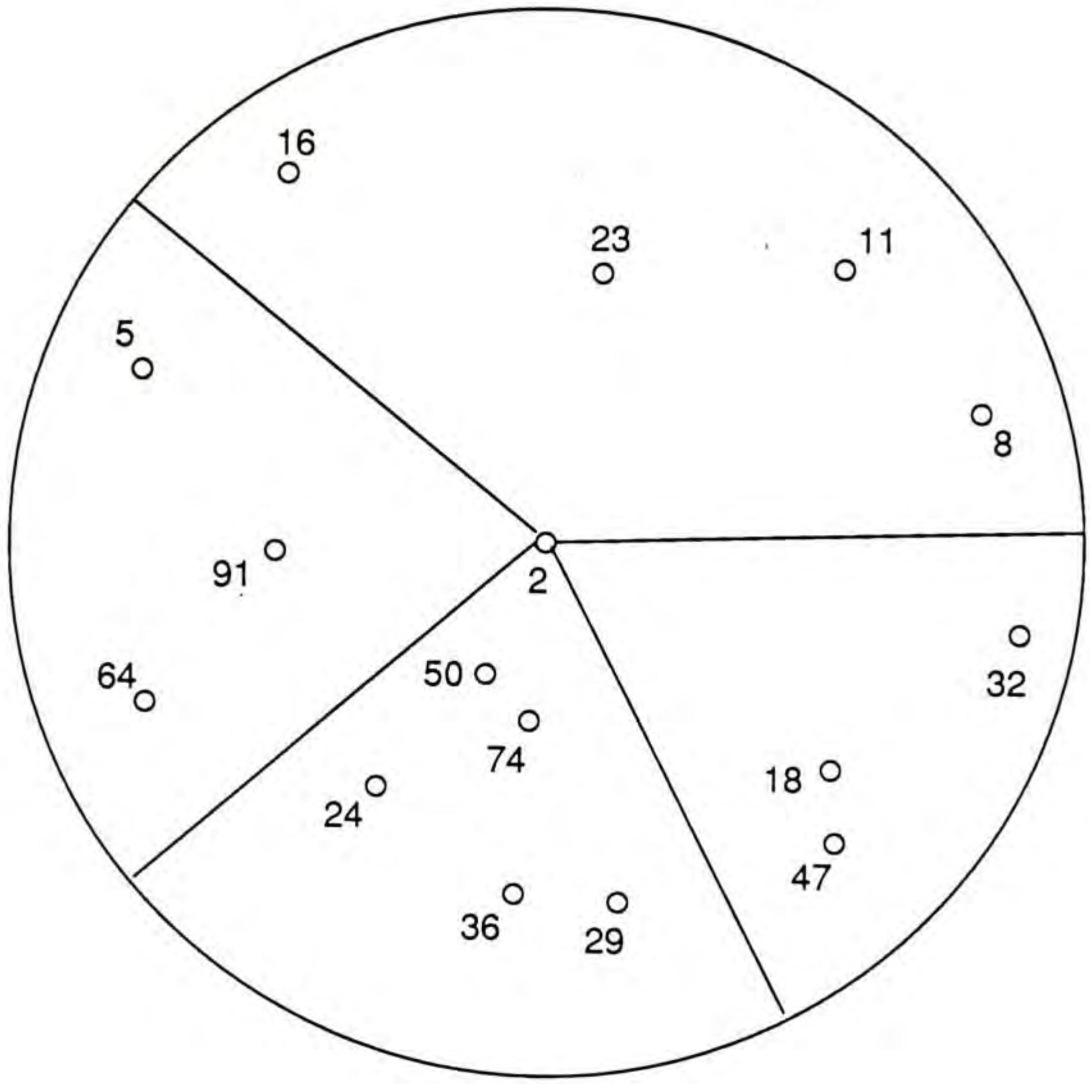


Fig.2.9 Orientation and broadcasting angles of the anteenas

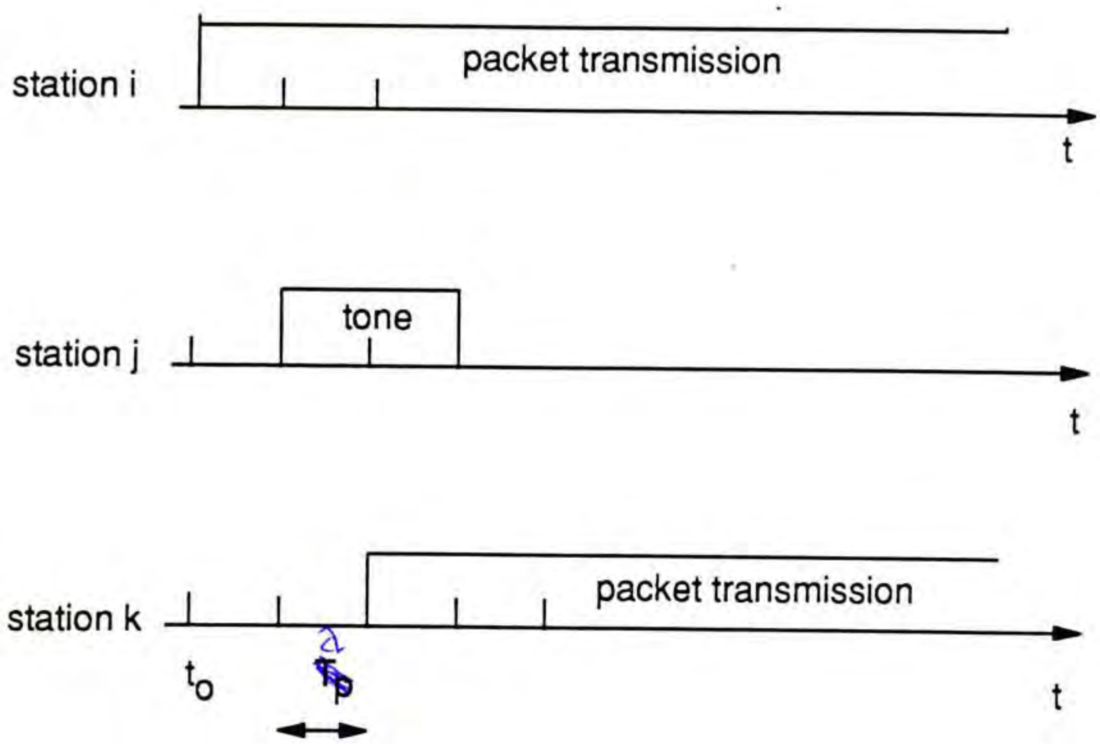


Fig.2.10 Determination of T_1

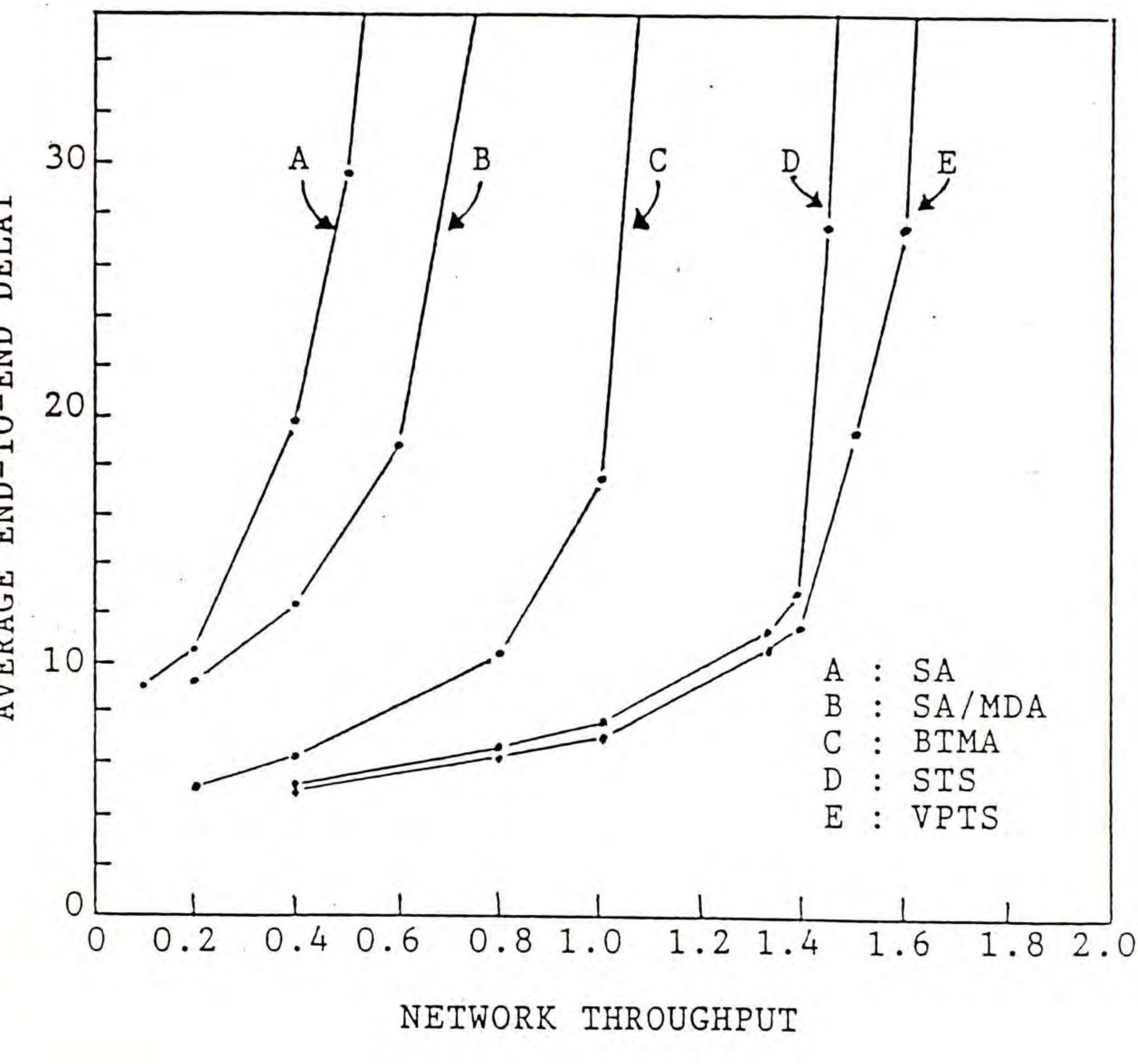


Fig.2.11 Delay vs throughput for a 80 node network

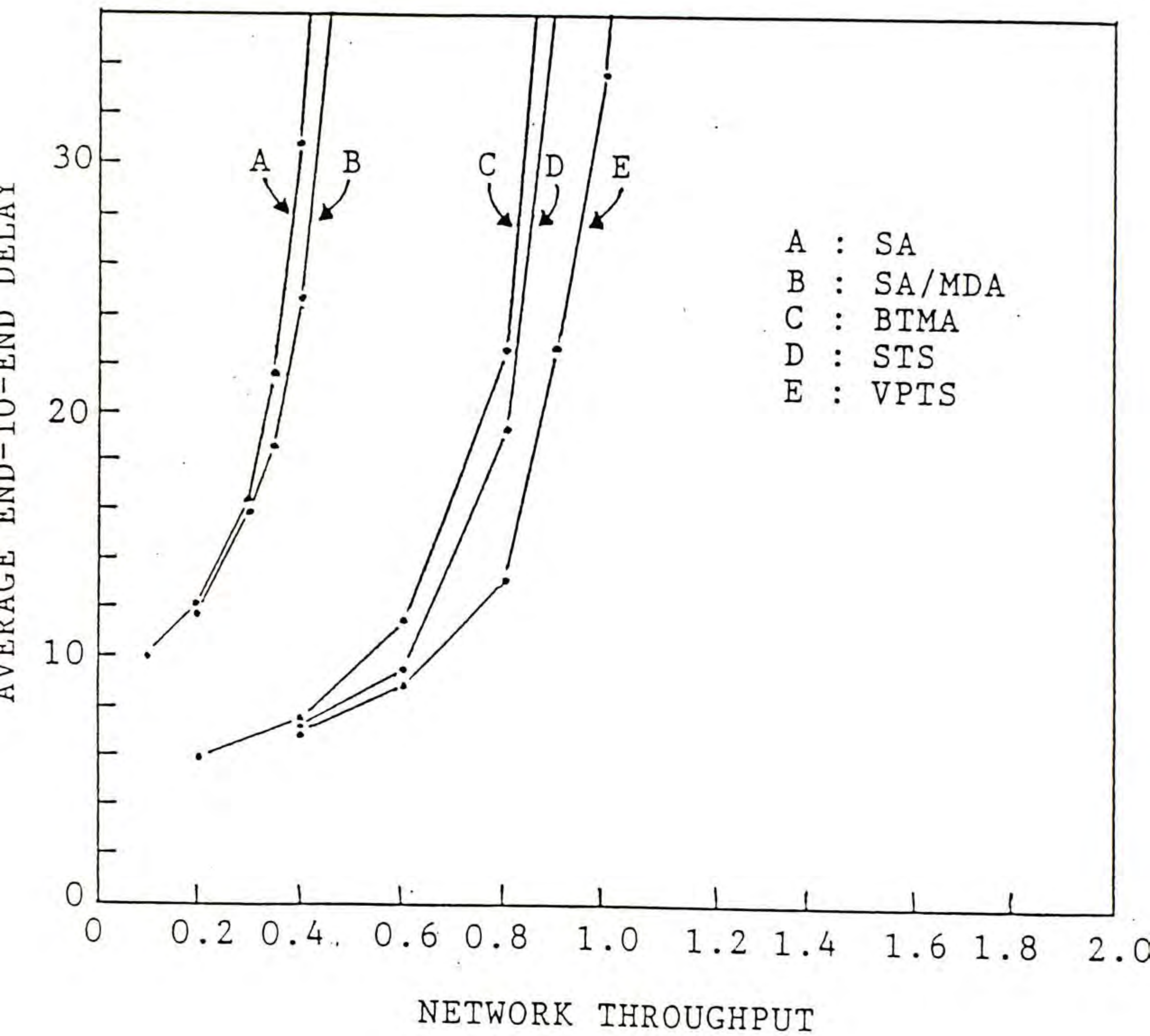


Fig.2.12 Delay vs throughput for a 40 node network

CHAPTER 3

THE CODED TONE SENSE PROTOCOL

3.1 INTRODUCTION

It is well known that the CSMA protocol can give a higher throughput than the ALOHA protocol in a centralized PRN. But its performance degrades in a multihop network environment. This is mainly due to the hidden station problem which could be solved by the use of a busy tone. All the above protocols are primarily designed for use with conventional radio signals. If there is overlap of transmissions from different stations, all the packets involved would be destroyed. In spread spectrum techniques, the radio signal is encoded using pseudo-random sequences. The spreading sequences permit the receivers to distinguish one spread-spectrum transmission from another. The use of a CDMA protocol allows overlapping of transmissions by assigning a different code to each transmitted signal and influences the choice of channel access protocols in PRNs.

Brazio and Tobagi presented a model for the throughput analysis of multihop spread spectrum PRNs in [BRAZ 85]. The access protocols considered include nonpersistent CSMA, pure ALOHA, conservative BTMA and Destination Code Sensing Multiple Access (DCSMA). Numerical results are only shown for some simple topologies with 3 to 4 nodes. In DCSMA, the source station monitors the channel for the transmission using the destination code of its packet prior to the transmission of its packet. This protocol is the same as the Receiver-base CSMA protocol [GERA 88] where the CSMA protocol is embedded on a single hop spread-spectrum PRN with a unique spreading code allocated to each station for receiving packets.

Chen and Boorstyn presented an approximate throughput analysis of a CDMA protocol in multihop PRNs [CHEN 85]. The effect of connectivity on network throughput in the presence of noise was also investigated. In [SOUZ 85], the study was extended to include BTMA and Preamble Sense Multiple Access (PSMA). In PSMA, a station will receive a packet from a neighbour, whether addressed to it or not, only if the initial portion (preamble) of the packet is not interfered. A level of noise immunity for the CDMA protocol is defined. Thus in CDMA/n, an idle station can successfully receive a new packet if there are less than n transmissions in its neighbourhood. It was concluded that for random networks and uniform end-to-end traffic, the protocols can be ranked in order of performance as CDMA/∞,..., CDMA/2, BTMA, CSMA, PSMA and CDMA/1 (ALOHA).

In multihop PRNs, a large number of stations requires a large number of spreading codes and hence a larger channel bandwidth. Moreover as the performance of the CSMA protocol degrades in a multihop environment, a more suitable protocol is needed to make good use of spread-spectrum techniques. Since the spreading codes assigned to the stations need to be unique only to its neighbours, the codes could be reused by stations which are farther apart. In this chapter, we first propose a code assignment algorithm based on the code reuse property. This algorithm could reduce the number of codes required to about 20% of the total number of stations in the network. Further reducing the number of codes is found to cause little throughput degradation. We then design a new protocol termed Coded Tone Sense (CTS) for using these codes in Section 3.3. Code assignment examples and simulation results are presented in Section 3.4.

3.2 SYSTEM MODEL AND CODE ASSIGNMENT ALGORITHM

Let there be N stations in a packet radio network and let their locations be fixed. Each station is assigned a code and a station-number. The station-number is globally unique, but the code is unique only in each station's neighbourhood. Let the transmission range be R for all stations. Each station has only one receiver and one transmitter and all stations use the same frequency band for transmitting packets. Stations therefore cannot transmit and receive at the same time.

Each station is assigned a code for identifying itself from its neighbours. Since the codes are local in nature, beyond a certain range, which we call it the local-range for convenience, they can be reused. The size of the local-range depends on R and the distribution of neighbouring stations.

Each station is first assigned a code to distinguish itself from the other stations within its local-range. This code assignment problem is just the same as the tone assignment problem in Chapter 2. Hence the same algorithm in Section 2.5.1 can be used for assigning codes to stations and is repeated as follows:

(1) j := 1.

(2) Select an unassigned station and denote it as S_o.

(3) Assign code A_j to S_o.

(4) S := S_o.

(5) Mark all the stations in the local-range of station S.

(6) If all unassigned stations are marked, go to (9).

(7) Assign code A_j to one of the unmarked stations S'.

(8) S := S' ; go to (5).

(9) If all stations are assigned, stop. Otherwise unmark all unassigned stations.

(10) j := j + 1 ; go to (2).

We denote the total number of codes required using the above algorithm as K_1 and the initial code assigned to station j as code A_j. Note that K_1 codes are needed to avoid code collisions. But in PRNs, collision of packets due to time conflict is common. Therefore, if code collision can be tolerated, the total number of codes K_1 can be reduced to save bandwidth. We shall show in Section 3.4 that when the number of codes is reduced to a small fraction of K_1 only a small throughput degradation is observed. The criteria of sharing codes depends on the station distribution and the traffic on the network. Here we choose, for simplicity, to allocate codes so that the number of stations sharing a code is as even as possible. Let K_2 ($\leq K_1$) be the desired number of codes and C_j be the final code assigned to station j. Then

$$C_j = A_j \mod K_2 \quad .$$

3.3 PROTOCOL DESCRIPTION

To avoid being interfered by the neighbouring stations, a station will broadcast a busy tone during its packet reception. The receiving station will stop the busy tone when collision occurs. The transmitting station can detect the collision by monitoring the busy tone of the destination. When the number of codes used in the network is K_2, the number of different busy tones required is also K_2. Each code has a corresponding busy tone. The K_2 busy tones can be K_2 different frequencies or just another K_2 spreading codes. Each station keeps a Code Table to record the codes and tones of all its neighbours.

3.3.1 Transmission Protocol

(1) Find the code C_j and tone B_j of the receiving station from the Code Table and encodes the data packet using code C_j.

(2) Sense tone B_j. If B_j is detected, go to step (2) after a random delay. If B_j is not detected, transmit the packet immediately.

(3) During the packet transmission, if B_j is not detected in the time-out period or B_j terminates during the packet transmission, stop the transmission immediately, wait for a random delay and go to step (2).

3.3.2 Reception Protocol

(1) When an incoming packet is detected in the station's assigned code C_j (i.e. after receiving the packet header), broadcast its assigned tone B_j immediately.

(2) When the station detects a collision or error while receiving a packet, stop the busy tone immediately.

3.4 SIMULATION RESULTS

Four network samples are generated on which the performance of various protocols are compared. The stations in the network are randomly located within a 20 km x 20 km square region. The transmission range is 4 km. The packet generation rates are the same for all stations and the packet destinations are equally probable for all stations, excluding the source station. Let the packets be of fixed length and let the arrivals to each station be a Poisson process. Minimum hop routing rule is used. The characteristics of the networks generated are summarized as follows:

Network parameters \ Cases	1	2	3	4
No. of stations N	80	80	40	40
ave. no. of neighbours per station	8.31	8.38	4.05	5.45
max. no. of neighbours per station	15	12	7	11
No. of codes K_1	21	15	9	14

Using the code assignment algorithm, the number of codes K_1 required without interference is reduced to 20% - 35% of the total number of stations N. We denote the Coded Tone Sense protocol with n codes as CTS/n and set n=5 for the 80 station networks and n=3 for the 40 station networks in our examples. Note that n is the desired number of codes K_2 and CTS/1 is just the BTMA protocol. Using the code assignment algorithm on the Slotted ALOHA protocol, we have the Coded Slotted ALOHA (CSA/n) protocol. For cases 1 and 3, we also compared CSA/n with CTS/n.

The average end-to-end delay as a function of network throughput for cases 1 and 3 are plotted in Figures 3.1 and 3.2 respectively. The throughput-delay characteristics of cases 2 and 4 are similar and so are not shown. The maximum network throughput attained for the four station distributions (or the four cases) are obtained as follows:

Protocols \ Cases	1	2	3	4
SA	0.5	-	0.35	-
CSA/n	0.9	-	0.5	-
CSA/K_1	1.1	-	0.55	-
BTMA	1.1	1.1	0.8	0.8
CTS/n	1.9	2.0	1.0	1.0
CTS/K_1	2.0	2.2	1.1	1.2

The maximum network throughput of CTS/5 is found to be 73% to 82% higher than that of BTMA for the 80 station networks. For the 40 station networks, CTS/3 gives about 25% improvement. When the number of code groups is increased to K_1, there is only 5% to 10% further improvement for the 80 station networks. For the 40 station networks, the further improvement is from 10% to 20%. CTS/K_1 always have a smaller delay than CTS/n.

The maximum network throughput of CSA/n is found to be 80% higher than that of SA for case 1 and 42% higher than that of SA for case 3. When the number of code groups is increased to K_1 , there are 22% and 10% further improvements for cases 1 and 3 respectively.

It can be concluded that there is a performance improvement of using more codes when the stations are densely located. For the 80 station networks using only 5 codes, the network performance is almost the same as those using 21 codes.

3.5 CHAPTER SUMMARY

Using spread spectrum techniques in PRNs, overlapping of packet transmission is allowed by assigning a different code to each transmitted signal. We have designed an algorithm for assigning codes to the stations such that these codes can be reused beyond their interference range. This algorithm can reduce the number of spreading codes required to 20% - 35% of the total number of stations in the network.

Using the code assignment algorithm on Slotted ALOHA, the resulting CSA/n protocol can give 42% to 80% performance improvement over the SA protocol. We have also designed the Coded Tone Sense protocol which can further reduce the number of codes required. From simulation results, it was found that the CTS protocol has a much better performance than the BTMA protocol. For a 80 station network using only 5 codes, the maximum throughput of the CTS protocol is found to be 73% to 80% higher than that of the BTMA protocol.

It was found that the CSA and CTS protocols are particularly attractive for densely populated networks. For these networks only a few codes is sufficient to drive the throughput-delay performance very close to the case where each station has a unique code.

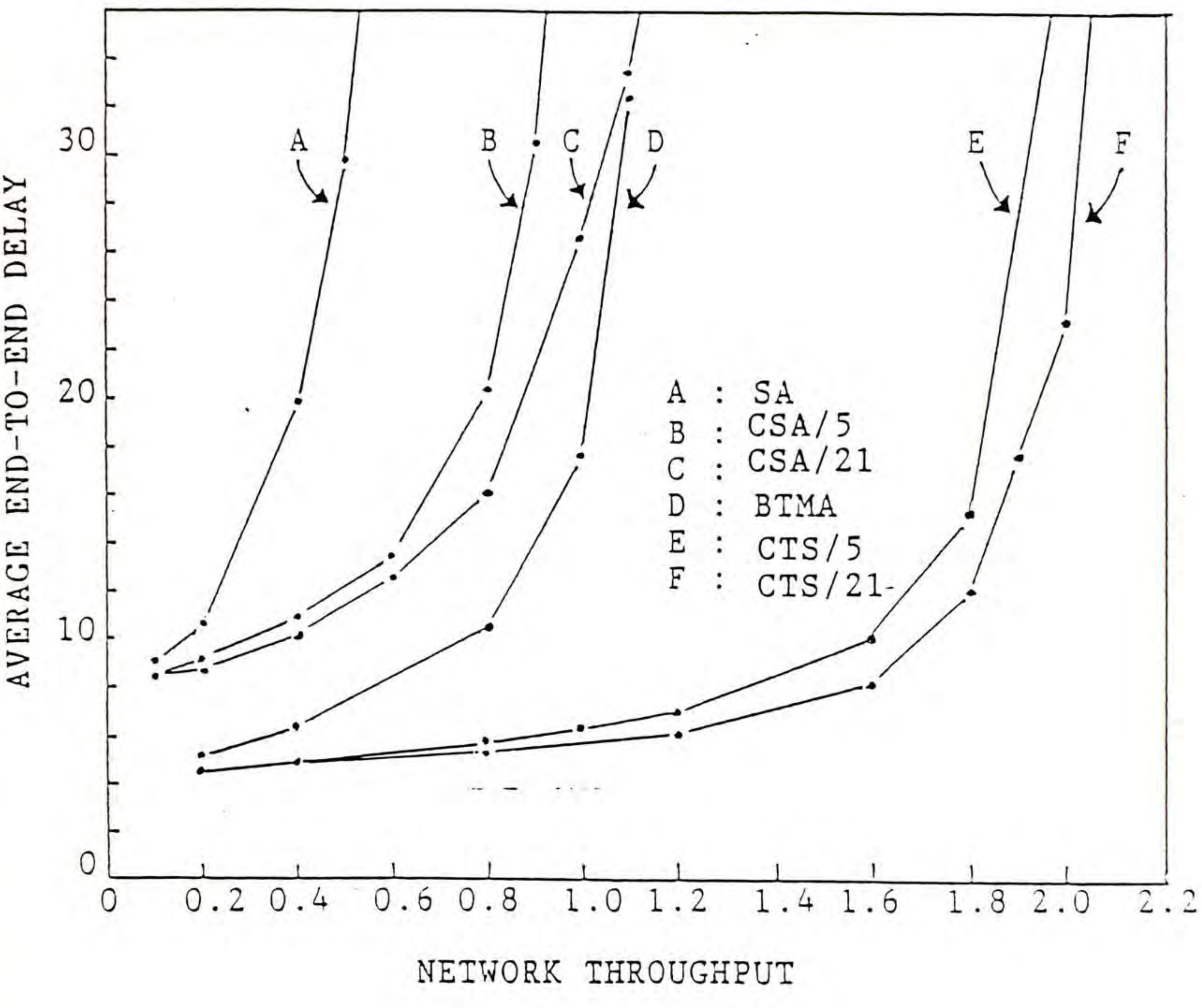


Fig.3.1 Delay vs throughput for a 80 node network

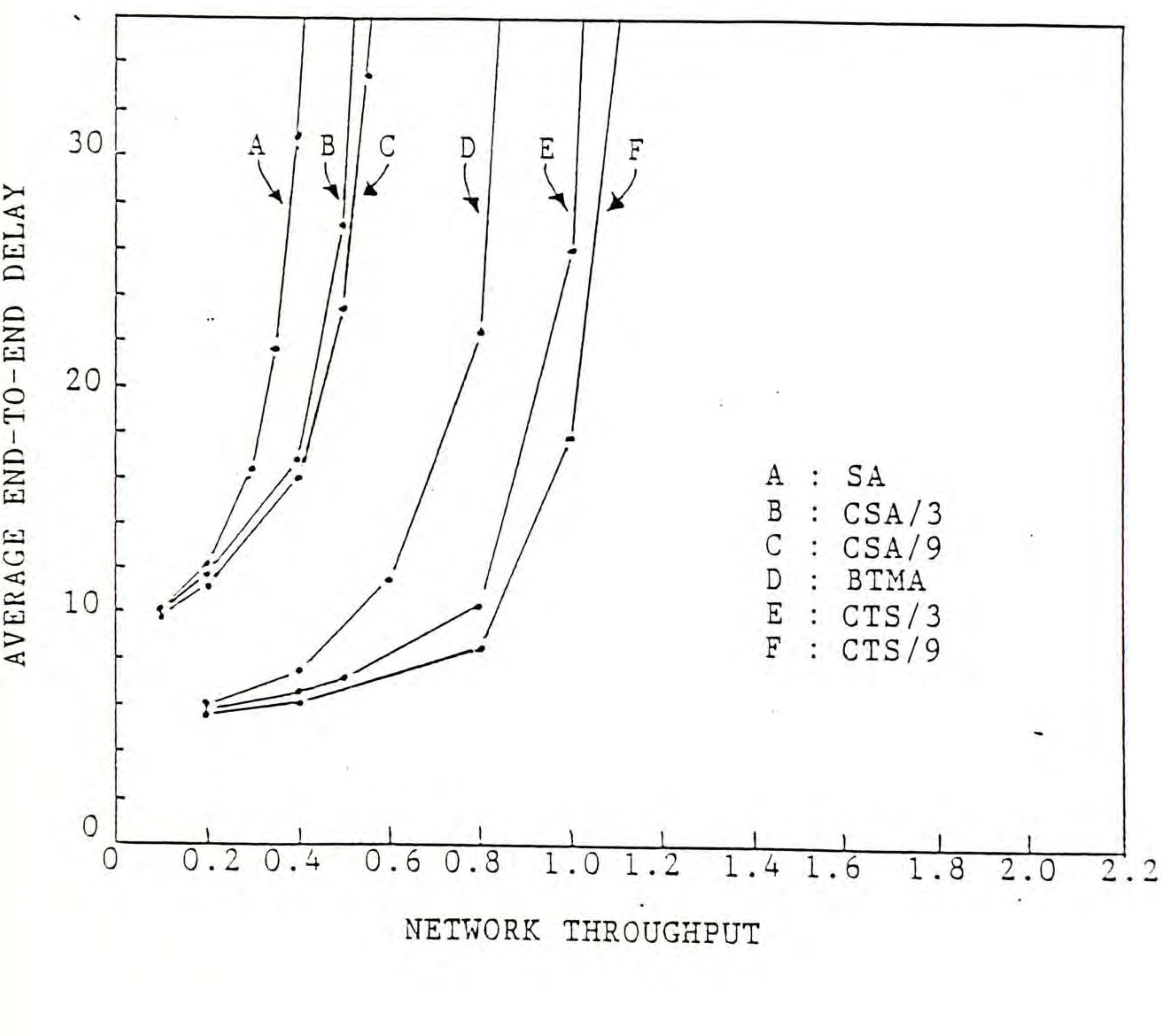


Fig.3.2 Delay vs throughput for a 40 node network

CHAPTER 4

AN EFFICIENT SPREADING CODE ASSIGNMENT ALGORITHM

4.1 INTRODUCTION

Spread spectrum signaling opens up a new dimension for protocol design and performance trade-off. When the receiver-based or transmitter-based spreading code protocols are used, a multihop PRN with a large number of stations requires a large number of codes and hence a large channel bandwidth. Since a spreading code assigned to a station needs to be unique only to its neighbours, the codes could be reused by the stations which are farther apart. The code assignment algorithm presented in Section 3.2 can reduce the number of codes required but does not minimize the number of codes needed.

It is important to find an efficient algorithm for assigning as few codes to the PRN stations as possible since the smaller the number of codes used the smaller the bandwidth needed. In this chapter we shall first transform the code assignment problem to the familiar graph coloring problem. This allows us to consider the possible use of the graph coloring algorithms for assigning codes in PRNs. We then design a heuristic code assignment algorithm making use of some special properties of PRNs. We also obtain a lower bound on the chromatic number, which in our case is the minimum number of codes required. Finally the performance of this new algorithm is assessed by making comparisons to the bound as well as to one of the best heuristics for graph coloring.

4.2 CODE ASSIGNMENT AND GRAPH COLORING

The coloring of a graph G means assigning colors to the vertices of G so that adjacent vertices have different colors [TUCK 84]. The chromatic number of a graph is defined as the minimum number of colors needed to color the graph. A variety of problems in computer science, operations research, and the design of experiments have been identified as the graph coloring problem. The problem of showing that any map can be 4-colored tantalized mathematicians for 100 years until a computer-assisted proof was obtained by Appel and Haken in 1976 [APPE 76]. This problem is a particularly interesting special case of the graph coloring problem.

There are numerous papers on graph coloring. It was shown that this problem is NP-complete in the sense of Karp [KARP 72]. Therefore to solve large size graph coloring problems, many heuristics are proposed. One particularly good heuristic is the Degree Saturation or Dsatur Algorithm [BREL 79]. In order to measure the performance of the heuristics, various methods have been proposed to estimate the chromatic number of graphs [MITC 76] [MCDI 79].

The code assignment problem in PRNs is as follows. Let there be N stations in a packet radio network and let their locations be fixed. Each station is assigned a code which is unique only in its neighbourhood. Let the transmission range be R for all stations. Since the codes are local in nature, beyond a certain range, which we call it the local-range, they can be reused. The local-range of a station is formed by the perimeters of the transmission ranges of the neighbours of that station.

To relate the code assignment problem to the graph coloring problem, the network structure must first be represented by a graph. The stations in a multihop PRN are treated as the vertices and an edge is formed between two vertices when the two stations are neighbours. The graph G obtained can be represented by the adjacent matrix $A = \{a_{ij}\}$ where

a_{ij} = 1 if station i and station j are neighbours

= 0 otherwise.

Since stations separated by two hops also cannot be assigned to the same code, edges between all vertices which are two hops apart are added to reflect this requirement. With that, the code assignment problem of graph G becomes the coloring problem of graph G' where G' has adjacent matrix $B = \{b_{ij}\}$ such that

b_{ij} = 1 if $\{a_{ij} = 1$ or ($a_{ik} = 1$ and $a_{kj} = 1$ for any k)$\}$

= 0 otherwise.

Here $b_{ij} = 1$ means station i and station j are within the interference range of each other and should have different codes. Figure 4.1 shows an example of transforming G to G'.

The coloring problem can be solved by the Branch and Bound algorithm. For the efficient execution of the Branch and Bound algorithm a very tight lower bound on the chromatic number is very desirable. This is also true when a heuristic is used. In general it is fairly difficult and time consuming to obtain a good lower bound on the chromatic number of a large graph [MCDI 79]. However for the transformed graph G' that corresponds to the code assignment problem, a very tight lower bound on the chromatic number is readily available. The following is a derivation of that bound.

In graph theory the degree of a vertex is defined as the number of edges incident on that vertex. In PRNs we refer the degree of a station as the number of stations within that station's local-range. Let D_{max} denote the maximum degree among all the stations in a PRN and let C denote this network's chromatic number (i.e. the minimum number of codes needed). Let $k_1, k_2,..., k_N$ be the number of neighbours of station 1, 2,..., N. Then $k = \max(k_1, k_2,..., k_N)$ is the size of the largest neighbouring group.

All stations within the local-range of a particular station say A cannot use the same code as A. But they are not precluded to share codes. Hence the chromatic number can be less than the maximum degree D_{max} of the graph. In other words D_{max} is not a lower bound of the chromatic number. On the other hand, all neighbours of station A must be within the local-range of each other. Hence they cannot use the same code. Therefore the number of codes needed must be no smaller than the size of the largest neighbouring group plus one (i.e. station A itself). Hence a lower bound on the chromatic number C, denoted as C_L, is

$$C_L = k + 1 \quad .$$

4.3 ALGORITHM DESCRIPTION

In designing the code assignment algorithm, the following criteria were followed:

(1) When building a K-coloring of a graph, we can ignore all vertices of degree less than K, since once the other vertices are colored, there will always be at least one color available for each of these vertices [TUCK 84].

(2) The lower bound of the chromatic number obtained in Section 4.2 is a good starting point.

(3) When there are more than one available codes, we choose the code which gives the minimum binding in assigning codes to other stations.

Let c be the total number of codes currently used. When a station, say station A, is assigned to code x, this assignment will affect the assignment freedom of the set of stations in station A's local-range and might increase the number of codes needed. Consider the example in Figure 4.2 where the number inside a circle is the code number already assigned to that station. Now the total number of codes currently used is 5. If code 1 is assigned to A, then B will have two choices from codes 3 and 5 and C ~~will~~ can only have ~~one choice from~~ code 3. Therefore the final total number of codes needed is still 5. However if code 3 is assigned to A, then B will still have two choices from codes 1 and code 5 but C will have no choice from the used codes. An additional new code 6 is thus required for C.

For the set of stations within station A's local-range let n_i, i = 0, 1,..., c-1, be the number of stations with i codes to choose from after A chooses code x and let D(A) be the degree of station A. The degree of freedom left for the assignment of the remaining codes can be measured by the *binding function F* defined as follows:

$$F(A,x) = \sum_{i=0}^{c-1} n_i D(A)^{-i}$$

It is easy to see that the smaller the F(A,x), the larger the freedom of assigning codes to the stations within A's local-range and hence the smaller the number of additional codes required.

The code assignment algorithm is as follows:

(1) Find the station with maximum number of neighbours and denote it as station S^*. Let the number of neighbours of station S^* be k. Assign code 1 to station S^*. Assign code 2 to ~~code~~ k + 1 to the neighbours of station S^*. Let c = k + 1 where c is the total number of codes currently used.

(2) Rank all the stations in order of decreasing degrees and denote them as $S_1, S_2, \ldots, S_N$ where N is the number of stations in the network.

(3) j := 1.

(4) If station S_j is assigned, go to (7).

(5) If $D(S_j) < c$, go to (7).

(6) Find the codes available to S_j. (Find the codes which are not used by the stations in the local-range of S_j.)

(a) If no code is available then c := c + 1 and assign code c to S_j.

(b) If only one code is available, assign that code to S_j.

(c) If more than one code are available, choose the code with the minimum F value and assign it to S_j.

(7) If j < N then j := j + 1 and go to (4).

(8) ~~To~~ assign codes to the remaining stations, repeat steps (3) to (7) but skip step (5). (This second loop is for assigning codes to stations with degree less than c.)

4.4 RESULTS AND DISCUSSION

Many random networks are generated to compare our code assignment algorithm with the Dsatur Algorithm [BREL 79] (similar to our algorithm but without steps (1) and (6c)) and the lower bound on chromatic number. The stations in the networks are randomly located within a 20km x

But (6c) is the mos
complexed computatio
If you make it

20km square region. The transmission range is 4 km. The number of stations in the network ~~ranges~~ ~~from 40 to~~ 160. For each ~~case~~ 50 random station distributions ~~are generated~~. Table 1 summarizes the assignment results.

Our code assignment algorithm runs very fast. For a 160 station networks the assignment is completed within half a minutes using a PC/AT. Out of a total of 200 cases, there are 90 cases where our code assignment algorithm requires fewer codes than the Dsatur Algorithm but only 6 cases where our code assignment algorithm requires more. When using our code assignment algorithm the average numbers of codes needed are 9.32, 16.32, 23.48 and 29.60 for the 40, 80, 120 and 160 station networks respectively whereas the Dsatur Algorithm requires 9.36, 16.74, 24.32 and 30.78 codes respectively.

When the network contains 40 stations our code assignment algorithm can reach the lower bound k+1 in 88% of the time. This means that our code assignment algorithm gives the optimal result at least 88% of the time. When the number of stations is increased to 160, there are still 30% of the cases reaching the lower bound and the average additional codes required is only 1.22. This therefore shows that both the code assignment algorithm and the bound on chromatic number are very good indeed.

4.5 CHAPTER SUMMARY

Using spread spectrum techniques in PRNs, overlapping of packet transmission is allowed by assigning a different code to each transmitted signal. A spreading code assigned to a station needs to be unique only to its neighbours, the codes therefore can be reused by the stations farther apart. We found that the code assignment problem could be transformed to the graph coloring problem and a very efficient algorithm for assigning codes to the stations in a PRN has been designed. A very tight lower bound on the number of codes needed has also been derived.

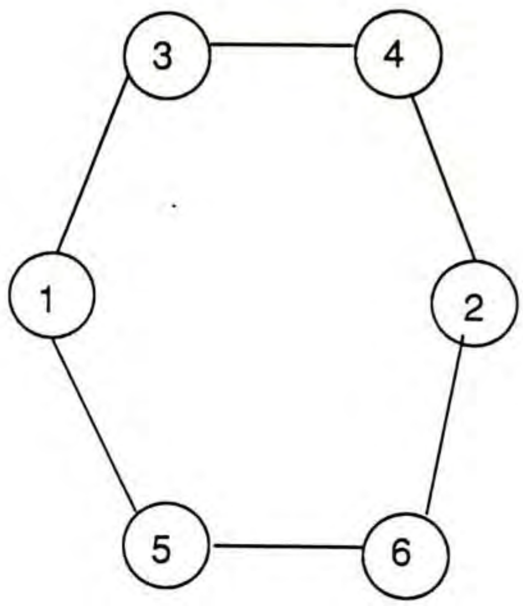
Network Parameters \ Network Type (1)	1	2	3	4
no. of stations N	40	80	120	160
max. size of neighbouring group k (2)	8.20	14.90	21.90	27.38
max. degree D_{max} (2)	15.02	33.28	53.26	71.04
average code size (2)	9.32	16.32	23.48	29.60
no. of cases reaching the lower bound	44	32	26	15
max. additional codes above the lower bound	1	2	2	4
no. of cases with smaller code size than Dsatur Algorithm	2	21	30	37
no. of cases with larger code size than Dsatur Algorithm	0	1	1	4
max. no. of codes saved compared to Dsatur Algorithm	1	2	3	4

Remark : (1) 50 sample networks were generated for each network type.

(2) averaged over the 50 samples.

Table 4.1 Code Assignment Results

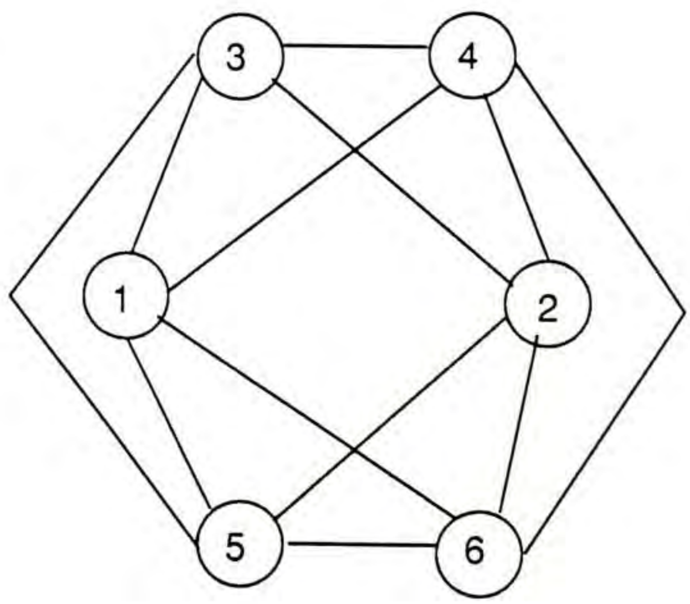
	1	2	3	4	5	6
1	0	0	1	0	1	0
2	0	0	0	1	0	1
3	1	0	0	0	0	1
4	0	1	0	0	1	0
5	1	0	0	1	0	0
6	0	1	1	0	0	0



The adjacent matrix of G

The graph G

	1	2	3	4	5	6
1	0	0	1	1	1	1
2	0	0	1	1	1	1
3	1	1	0	0	1	1
4	1	1	0	0	1	1
5	1	1	1	1	0	0
6	1	1	1	1	0	0



The adjacent matrix of G'

The graph G'

Fig.4.1 The graphs G, G' and their adjacent matrices

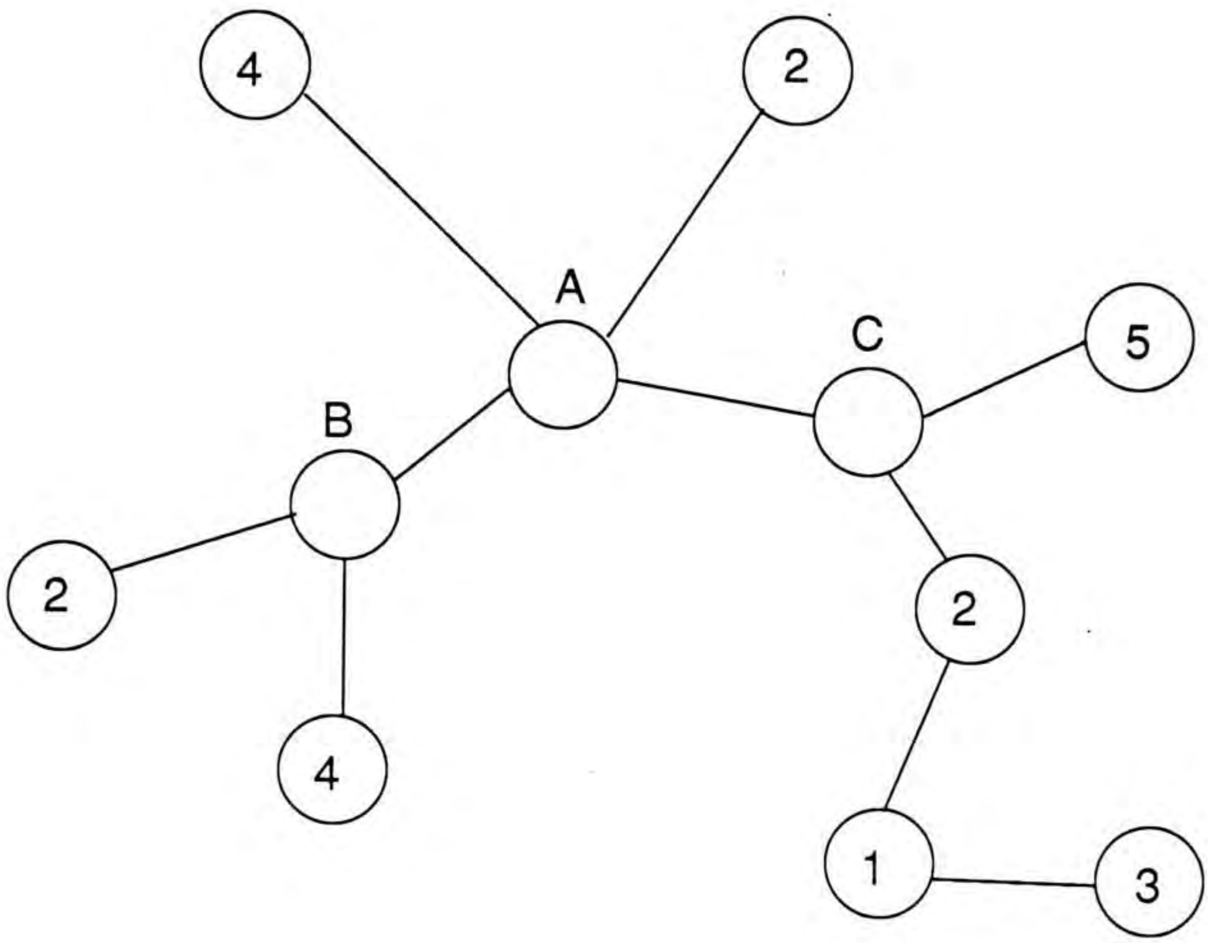


Fig.4.2 Restricted freedom of code assignment

CHAPTER 5

FAIR AND EFFICIENT TRANSMISSION SCHEDULING

5.2 INTRODUCTION

Transmission scheduling is one of the methods used to resolve the contention of the channel among neighbouring stations. This method amounts to assigning time slots to stations (or schedule transmissions) in such a way that stations transmitting in their assigned slots only will not encounter conflicts. In multihop PRNs, two stations can be scheduled to transmit at the same slot when they are far apart enough.

Previous studies on transmission scheduling in multihop PRNs include the work by Chlamtac and Kutten [CHLA 85] which showed that the problem of finding an optimal broadcasting protocol is NP-hard and proposed a polynomial time algorithm for generating Collision-Free Broadcast Spanning Trees. By broadcasting, they meant propagating a packet to *all* stations in the network. Thus using the broadcast spanning tree, a broadcast message can find its way to all destinations. Channel reuse, however, is prohibited for all nonbroadcast transmissions. Passive acknowledgement was used and both frequency-division and time-division approaches to protocol implementation were suggested. Later, this method was generalized in [CHLA 87] to allow spatial reuse of the channel among nonbroadcast transmissions. Centralized and distributed algorithms for constructing broadcast trees were derived. A station executing the broadcasting protocol must refrain from transmission in those slots assigned to its "father" (in the tree) and its neighbours' fathers. Following the reception of the broadcast message from its father, the station will forward the message in the next assigned time slot. The case of multiple broadcasting sources, however, was not considered.

Ephremides and Truong [EPHR 90] showed that the problem of scheduling "packet broadcast" in a multihop PRN with optimum throughput is NP-complete and proposed a scheduling algorithm assuming that each station has up to two-hop connectivity information. In their context, packet broadcast means packet transmission to the neighbouring stations or local broadcast. To differentiate this with global broadcast we shall refer local packet broadcast simply as *packet transmission* and local broadcast schedules as *transmission schedules*. In [EPHR 90], the number of slots in a schedule cycle is chosen to be the same as the number of stations in the network. Using the scheduling algorithm a skeleton TDMA schedule is first formed. The skeleton schedule is then exchanged between neighbours. The slots that are not reserved or not blocked are assigned to the eligible stations. The percentage of slots assigned for transmission was provided as a performance measure. They observed that their algorithm produces schedules that are not very fair.

Ramaswami and Parhi [RAMA 89] showed that the problem of forming a minimum length schedule is NP-complete and combined the features in [CHLA 87] and [EPHR 90] to construct transmission schedules using only one-hop connectivity information. The centralized version of the algorithm uses a sequential graph coloring heuristic for constructing schedules. The distributed version of the algorithm uses a circulating token for coordinating the stations in the network.

When designing scheduling algorithms, besides maximizing throughput and ensuring fairness of channel access, scheduling delays of stations should also be made as small as possible. This could be achieved by minimizing the cycle length of the transmission schedule and distributing the slots more evenly among the stations. In this chapter, we first design a new scheduling algorithm for multihop PRNs that give minimum cycle length schedule with close to maximum network throughput. In addition, the schedule also gives each station a fair allocation of transmission capacity. We then derive a set of four performance measures for scheduling algorithms, namely, (1) the schedule cycle length, (2) the scheduling delays under light and heavy traffic conditions, (3) the minimum transmission capacity and (4) the normalized network capacity. These measures should give more comprehensive comparisons of scheduling algorithms. Finally, extensive case studies show that the new algorithm always gives the shortest schedule cycle, the smallest mean scheduling delay and the largest minimum transmission capacity when compared to the algorithms in [EPHR 90] and [RAMA 89]. It also gives the smallest difference between the normalized network capacity and minimum transmission capacity. This means that the new algorithm can more fairly allocate transmission capacity to the stations. The normalized network capacities obtained by the three algorithms are almost identical, indicating that all three algorithms are efficient.

5.3 THE SCHEDULING PROBLEM

Let there be N stations in a multihop packet radio network and let their locations be fixed. Let the transmission range be R for all stations and let all stations within a station's transmission range be called the neighbours of that station. Let all stations use the same frequency band for transmission. Stations therefore cannot transmit and receive at the same time. Let the slot duration be equal to the packet length plus the maximum round trip delay. Stations are allowed to transmit only at the beginning of a slot.

There are two kinds of conflict in PRNs. A *primary conflict* occurs when a station receives two or more transmissions destinated to it simultaneously. A *secondary conflict* occurs when a station receiving a particular transmission is within the range of another transmission for other stations. Secondary conflicts are tolerated in the case of spread spectrum signaling. In this chapter we only consider the case where both primary and secondary conflicts are not tolerated. Therefore, two stations can transmit in the same slot without conflicts only when they are more than two hops away from each other. Hence the one-hop and two-hop neighbours of a station should be scheduled to transmit in different slots. The slot assignment pattern is repeated after a schedule cycle, which is the minimum interval of time in which stations are assigned at least one slot for transmission. A schedule is said to be fair if the number of slots assigned to each station is evenly distributed.

The scheduling problem to be considered is as follows: under the restriction of no interference among one and two-hop neighbours, assign as many stations as possible to transmit such that the schedule formed should be as fair as possible and of minimum cycle length.

The problem of forming a minimum cycle length schedule can be transformed to the two-hop graph coloring problem [RAMA 89]. Since the scheduling problem is a special case of graph coloring, special property of PRNs can be used to modify existing coloring algorithm to obtain better results. In Chapter 4 we showed that the code assignment problem in multihop spread spectrum PRNs can also be transformed to the two-hop coloring problem. A very tight lower bound on the number of colors required was found to be equal to the maximum number of neighbouring stations plus one. We also designed a very efficient code assignment algorithm which only requires minimum number of colors. This algorithm is used in the first half of our scheduling algorithm to obtain minimum cycle length schedules.

5.4 THE SCHEDULING ALGORITHM

By using the algorithm in Section 4.3, a color, which is unique among other stations within two-hop distance, is assigned to each station. Stations belonging to the same color group can be assigned to the same transmission slot without conflicts. Therefore the first stage of our scheduling algorithm is similar to that in Section 4.3. During the first stage, a minimum length schedule cycle is formed. Each station is first assigned to one slot. The assignment should leave the greatest freedom for assigning additional stations to the same slot in the second stage.

In the second stage, the schedule table is examined to assign additional stations to the slots. We define the total number of one and two-hop neighbours of a station as the *degree* of that station. Stations with larger degree have more restriction in chosing additional transmission slots. These stations should be examined first to increase their chances of getting additional slots. In order to produce a fairer schedule the stations are examined in decreasing order of degree and they are assigned to only one additional slot in each round of the assignment. In the following scheduling algorithm, steps (1) to (7) constitute the first stage while steps (8) and (9) constitute the second stage.

(1) Find the station with the maximum number of one-hop neighbours. Assign different colors (starting from 1) to that station and its neighbours.

(2) Rank the stations in decreasing order of degree.

(3) Choose the next uncolored station according to the degree ranking.

(4) If the degree of the chosen station is greater than or equal to the number of colors currently used, assign the lowest numbered color to the chosen station. (This assignment should leave the greatest freedom in assigning colors to the remaining stations using the binding function defined in Section 4.3.)

(5) Return to (3) until all stations have been examined once.

(6) Assign the available colors to the remaining uncolored stations.

(7) Schedule the stations belonging to color group i to transmit in slot i. The schedule cycle should repeat after a period of L slots where L is the total number of color groups.

(8) Choose the next (in order of rank, modulo N) station with an unassigned slot, and denote it as station X. Stop when no such station is found (i.e. all slots have been assigned).

(9) Find the first available slot of station X that has not been assigned to the one-hop and two-hop neighbours of station X and assign that slot to station X. Return to (8).

Example: Figure 5.1 shows a 15 node network. An edge between two nodes indicates that the nodes are within the transmission range of each other. Table 5.1 shows the transmission schedule. The length of the schedule cycle is 8 slots which is the minimum since station J has 7 neighbours. In the table a letter T in the x row and y column indicates that station y can transmit in slot x. Slots marked with T* indicate that these are additional transmission slots obtained in the second stage of the algorithm.

5.5 PERFORMANCE ANALYSIS

For the following analysis scheduling delay is defined as the waiting time until the next transmission slot and network capacity is defined as the maximum attainable throughput of the whole network. For simplicity, only the mean scheduling delays under heavy and light traffic conditions are derived.

Let L be the schedule cycle length and T_k be the number of transmission slots assigned to a particular station, say, station k. For station k, let $S_1, S_2, \ldots, S_{T_k}$ be the slot numbers assigned for transmission where $1 \leq S_1 < S_2 < \ldots < S_{T_k} \leq L$.

Under heavy traffic conditions, we assume there are always packets waiting for transmission at each assigned slot. Thus after the transmission of a packet at slot S_i, another packet will reach the head of the transmission queue at slot $S_i + 1$ and will be transmitted at slot S_{i+1}. Therefore the scheduling delay d_i of this packet is

$$d_i = \begin{cases} S_{i+1} - (S_i + 1) & i = 1, 2, \ldots, T_k - 1 \\ S_1 - (S_i + 1) + L & i = T_k. \end{cases} \tag{5.1}$$

Let $D_1(k)$ be the mean scheduling delay of station k *under heavy traffic conditions* then, using (5.1) we have,

$$\begin{aligned} D_1(k) &= \frac{1}{T_k} \sum_{i=1}^{T_k} d_i \\ &= \frac{1}{T_k} \left(\sum_{i=1}^{T_k - 1} S_{i+1} - \sum_{i=1}^{T_k - 1} S_i - \sum_{i=1}^{T_k - 1} 1 + S_1 - S_{T_k} - 1 + L \right) \\ &= \frac{L - T_k}{T_k}. \end{aligned} \tag{5.2}$$

Let $D_{1,av}$ and $D_{1,\max}$ be defined as the average and the maximum of $D_1(1), D_1(2), \ldots, D_1(N)$. In other words,

$$D_{1,av} = \frac{1}{N} \sum_{k=1}^{N} D_1(k)$$

$$D_{1,\max} = \max\{D_1(1), D_1(2), \ldots, D_1(N)\}$$

These two measures are used to compare the scheduling delay performance for various scheduling algorithms on a given network under heavy traffic conditions.

Under light traffic conditions, a new arrival packet always becomes the head of the transmission queue. We partition the schedule cycle of station k into T_k intervals where each interval is consist of an assigned slot and the following consecutive idle slots. Let m_i be the length (in slots) of the i-th interval then

$$m_i = \begin{cases} S_{i+1} - S_i & i = 1, 2, \ldots, T_k - 1 \\ S_1 - S_{T_k} + L & i = T_k. \end{cases} \tag{5.3}$$

Let packets be equally likely to arrive in any of the L slots in a cycle and let e_j be the scheduling delay of the packet arrives at slot j. Then, the mean scheduling delay of station k *under light traffic conditions*, denoted as $D_2(k)$ is

$$\begin{aligned} D_2(k) &= \frac{1}{L} \sum_{j=1}^{L} e_j \\ &= \frac{1}{L} \sum_{i=1}^{T_k} (\text{sum of scheduling delays in the i-th interval}) \\ &= \frac{1}{L} \sum_{i=1}^{T_k} (0 + 1 + 2 + \ldots + m_i - 1) \\ &= \frac{1}{L} \sum_{i=1}^{T_k} \frac{(m_i - 1)m_i}{2} \end{aligned} \tag{5.4}$$

We define $D_{2,av}$ and $D_{2,\max}$ to be the same as $D_{1,av}$ and $D_{1,\max}$ but under light traffic conditions.

Now we proceed to find the network throughput which is the sum of the throughputs of all stations. Let C_k, the transmission capacity of station k, be defined as

$$C_k = \frac{\text{number of transmission slots assigned to station k}}{\text{schedule cycle length}}$$

$$= \frac{T_k}{L} \tag{5.5}$$

Let the network be loaded with the same λ packets/slot in all stations. Then the throughput of station k is

$$S_k(\lambda) = \begin{cases} \lambda & if\, \lambda < C_k \\ C_k & if\, \lambda \geq C_k \end{cases} \tag{5.6}$$

and the normalized network throughput is

$$S(\lambda) = \frac{1}{N} \sum_{k=1}^{N} S_k(\lambda) \tag{5.7}$$

The normalized network capacity, defined as $sup[\, S(\lambda)\,]$, is simply

$$S_{\max} = \frac{1}{N} \sum_{k=1}^{N} C_k \tag{5.8}$$

One measure of the fairness of capacity allocation is the transmission capacity of the station with minimum allocated capacity, denoted as $S_{\min}$. In other words,

$$S_{\min} = \min\{C_1, C_2, \ldots, C_N\} \tag{5.9}$$

A fair scheduling scheme should have $S_{\min}$ as large as possible, or as close to $S_{\max}$ as possible.

5.6 RESULTS AND DISCUSSION

We now turn to compare quantitatively the above scheduling algorithm with those proposed by Ephremides and Truong [EPHR 90] and Ramaswami and Parhi [RAMA 89]. For convenience we shall refer these three algorithms as H&Y, E&T and R&P algorithms respectively. In our comparisons we consider a square region of dimension 20km x 20km. A transmission range of 5 km is assumed. We consider four cases where the number of stations in the network are 25, 50, 75 and 100 respectively. For each case 50 random station distributions are generated. The network parameters, each averaged over the 50 sample station distributions, are summarized as follows:

	Case 1	Case 2	Case 3	Case 4
No. of stations	25	50	75	100
average degree	8.43	17.90	29.34	41.35
ave. no. of neighbours	4.14	7.92	11.78	15.93
max. no. of neighbours	7.38	14.06	20.22	26.86
cycle length (H&Y)	8.46	15.52	21.84	28.64
cycle length (R&P)	8.96	16.74	24.08	31.94
cycle length (E&T)	25	50	75	100

For each sample station distribution, three transmission schedules are computed by the three algorithms. Delay and throughput measures, $D_{1,av}, D_{1,\max}, D_{2,av}, D_{2,\max}, S_{\max}$ and $S_{\min}$, are then obtained for each of the schedules. Each of these measures are then averaged over the 50 samples under each case. Let us denote these "averaged" measures as $\overline{D}_{1,av}, \overline{D}_{1,\max}, \overline{D}_{2,av}, \overline{D}_{2,\max}, \overline{S}_{\max}$ and $\overline{S}_{\min}$.

Figure 5.2 (a) and (b) show the "averaged" scheduling delays $\overline{D}_{1,av}$ and $\overline{D}_{2,av}$ for the 4 cases. It is seen that the H&Y algorithm always has a smaller scheduling delay than R&P and E&T algorithms. For the 100 station case, the average scheduling delay with H&Y algorithm is reduced to 85% of that with R&P algorithm and 30% of that with E&T algorithm. Figure 5.3 (a) and (b) show the "averaged" scheduling delays $\overline{D}_{1,\max}$ and $\overline{D}_{2,\max}$ for the 4 cases and similar conclusions can be drawn on them. In all cases the scheduling delay under heavy traffic conditions is about twice of that under light traffic conditions. The scheduling delay is also seen to increase with the number of stations. This is obvious because as the same transmission bandwidth is shared by more stations the system is more congested.

Figure 5.4 (a) and (b) show the "averaged" normalized network capacity $\overline{S}_{max}$ and the "averaged" minimum transmission capacity $\overline{S}_{min}$ for the 4 cases. The normalized network capacities for the three algorithms are very close to each other in all cases. On the other hand, the H&Y algorithm always gives the largest value of $\overline{S}_{min}$ and hence the smallest difference between $\overline{S}_{max}$ and $\overline{S}_{min}$. Thus for the 100 station case the ratio of $\overline{S}_{max}$ to $\overline{S}_{min}$ is 1.39 in H&Y algorithm, whereas the ratios in R&P and E&T algorithms are 1.53 and 4.85 respectively.

5.7 CHAPTER SUMMARY

By properly scheduling transmission times, a lot of transmission conflicts can be avoided. This is particularly important in multihop packet radio networks where multiple simultaneous transmissions are allowed over non-interference regions. We have designed a very efficient scheduling algorithm based on the graph coloring technique that minimize the schedule cycle length and distribute the transmission slots more evenly among the stations. We have also derived (1) the schedule cycle length, (2) the scheduling delays under light and heavy traffic conditions, (3) the minimum transmission capacity and (4) the normalized network capacity as performance measures for scheduling algorithms.

Extensive case studies show that the new algorithm always gives the shortest schedule cycle, the smallest mean scheduling delay and the largest minimum transmission capacity when compared to the algorithms in [EPHR 90] and [RAMA 89]. It also gives the smallest difference between the normalized network capacity and minimum transmission capacity. The normalized network capacities obtained by the three algorithms are almost identical indicating that all three algorithms are efficient.

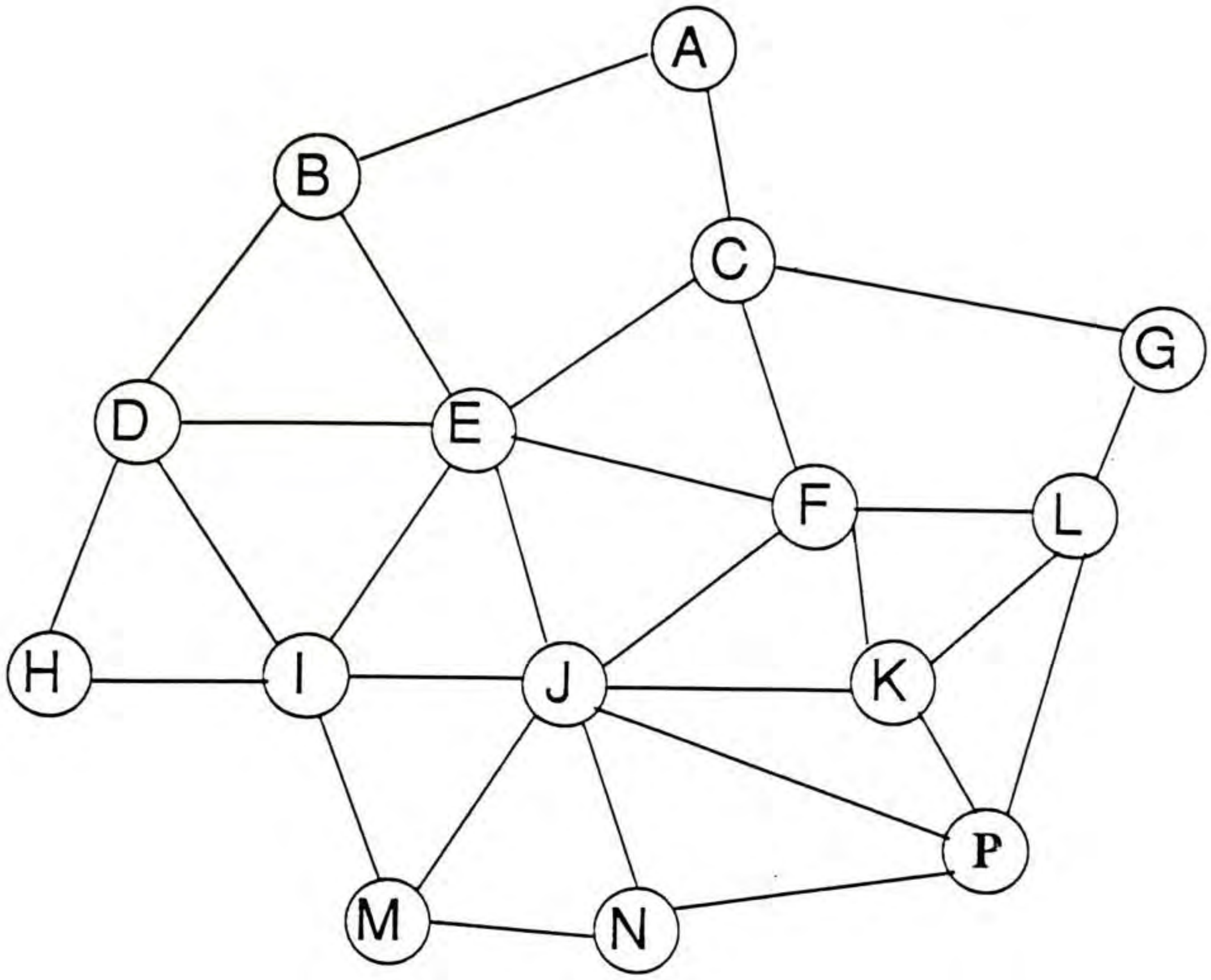
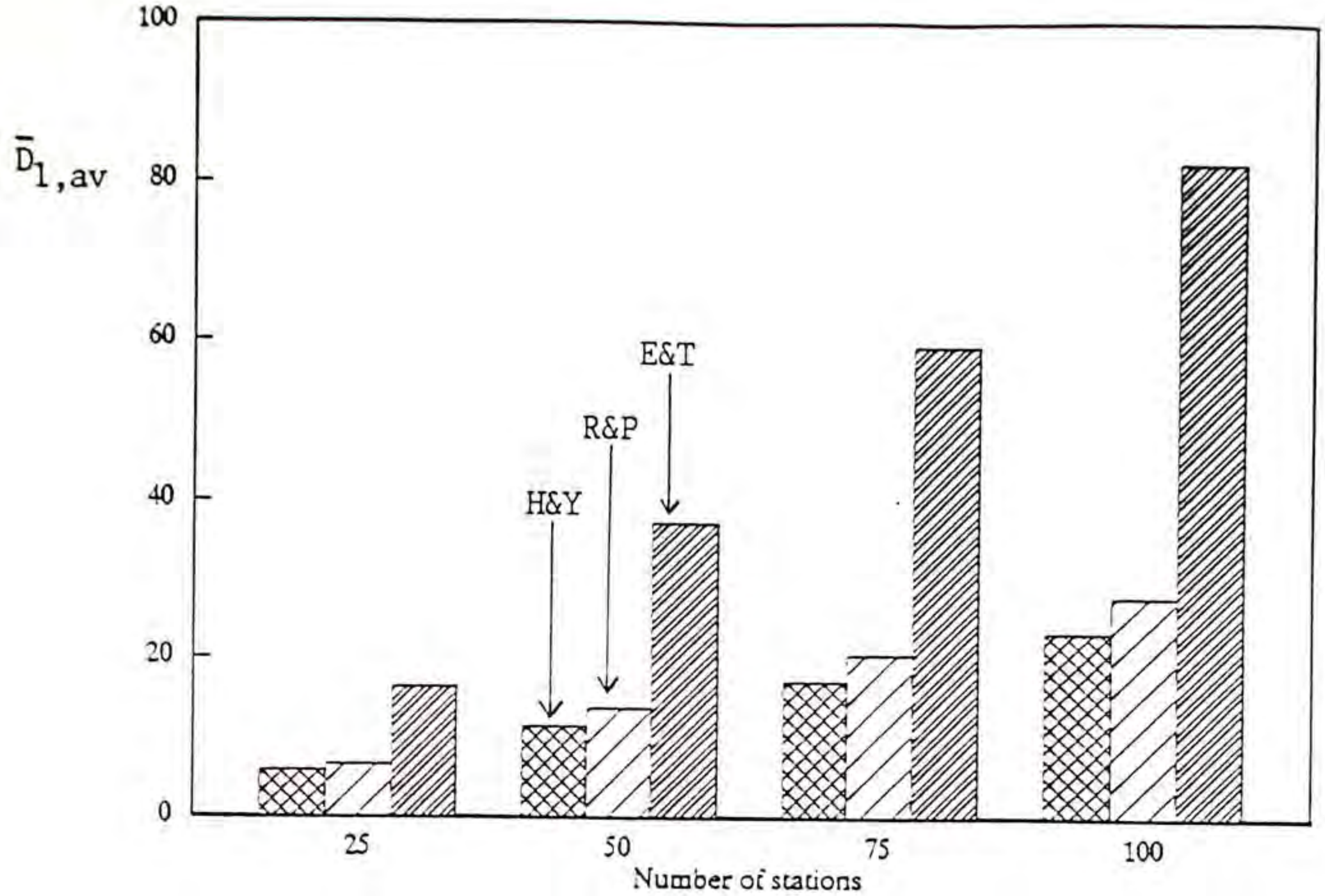


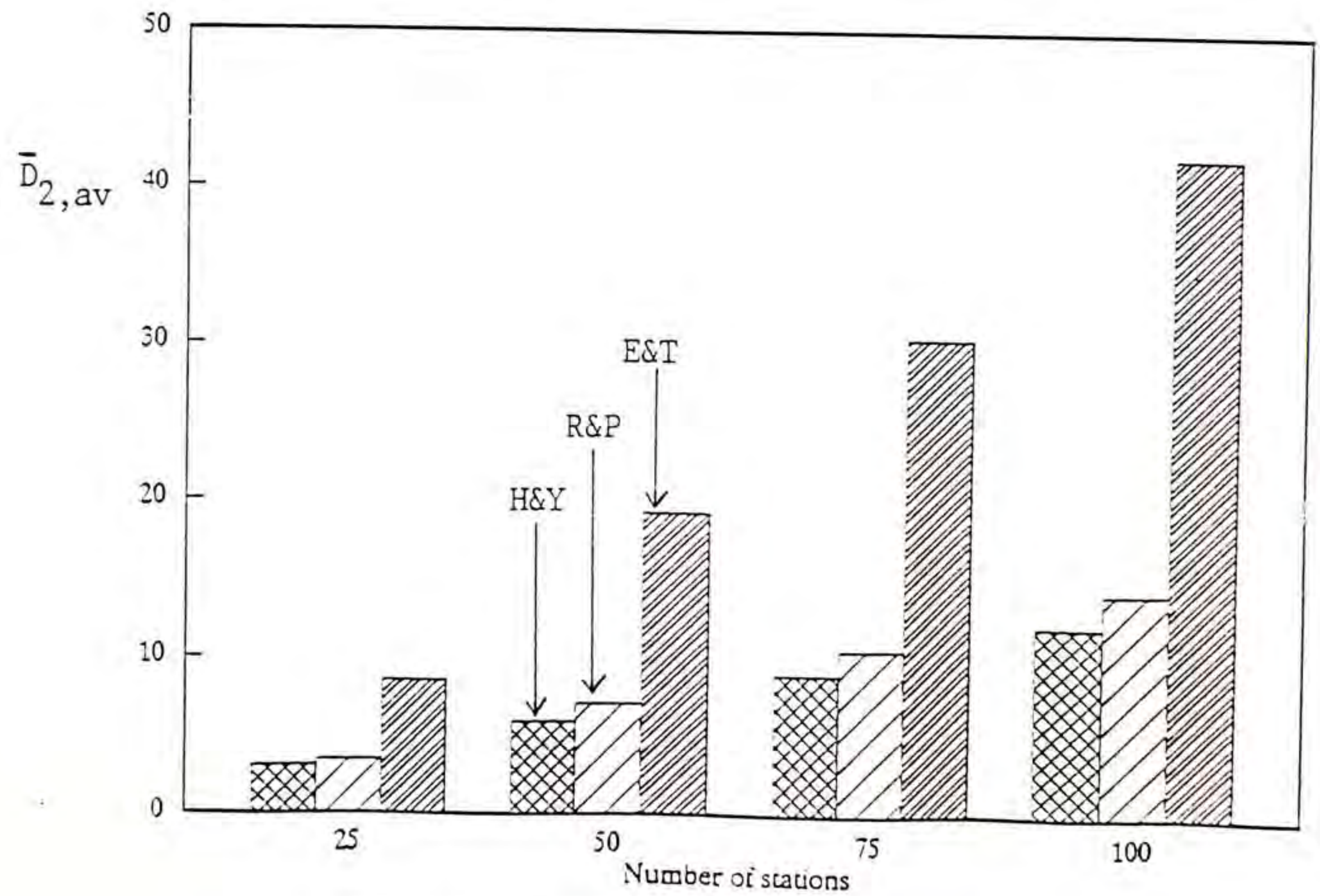
Fig.5.1 A sample network

Station ID / Slot Number	A	B	C	D	E	F	G	H	I	J	K	L	M	N	P
1	T									T					
2					T										
3						T		T							
4	T*								T			T			
5				T							T				
6		T					T						T		
7			T*					T*						T	
8			T					T*							T

Table 5.1 The transmission schedule

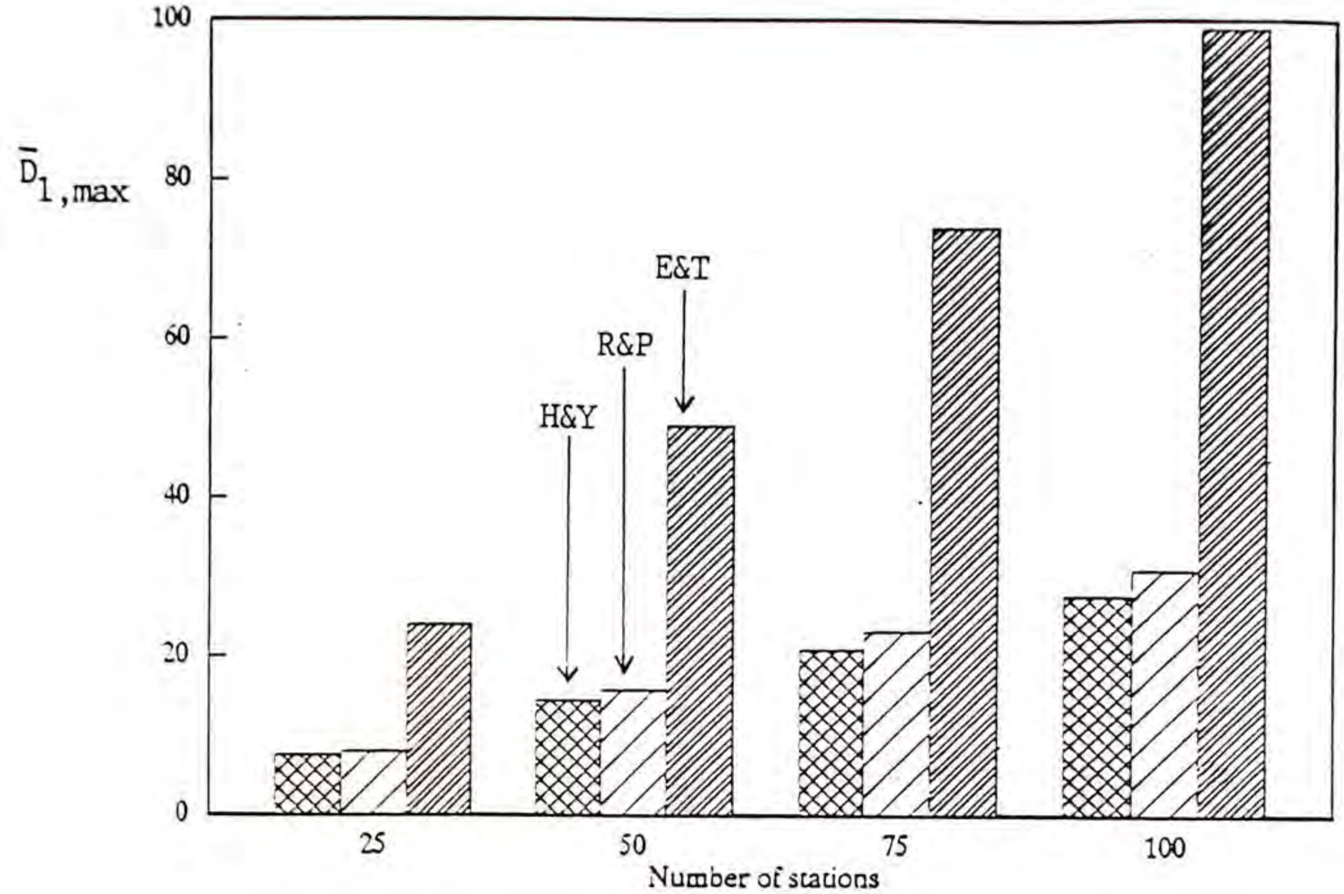


(a) under heavy traffic conditions

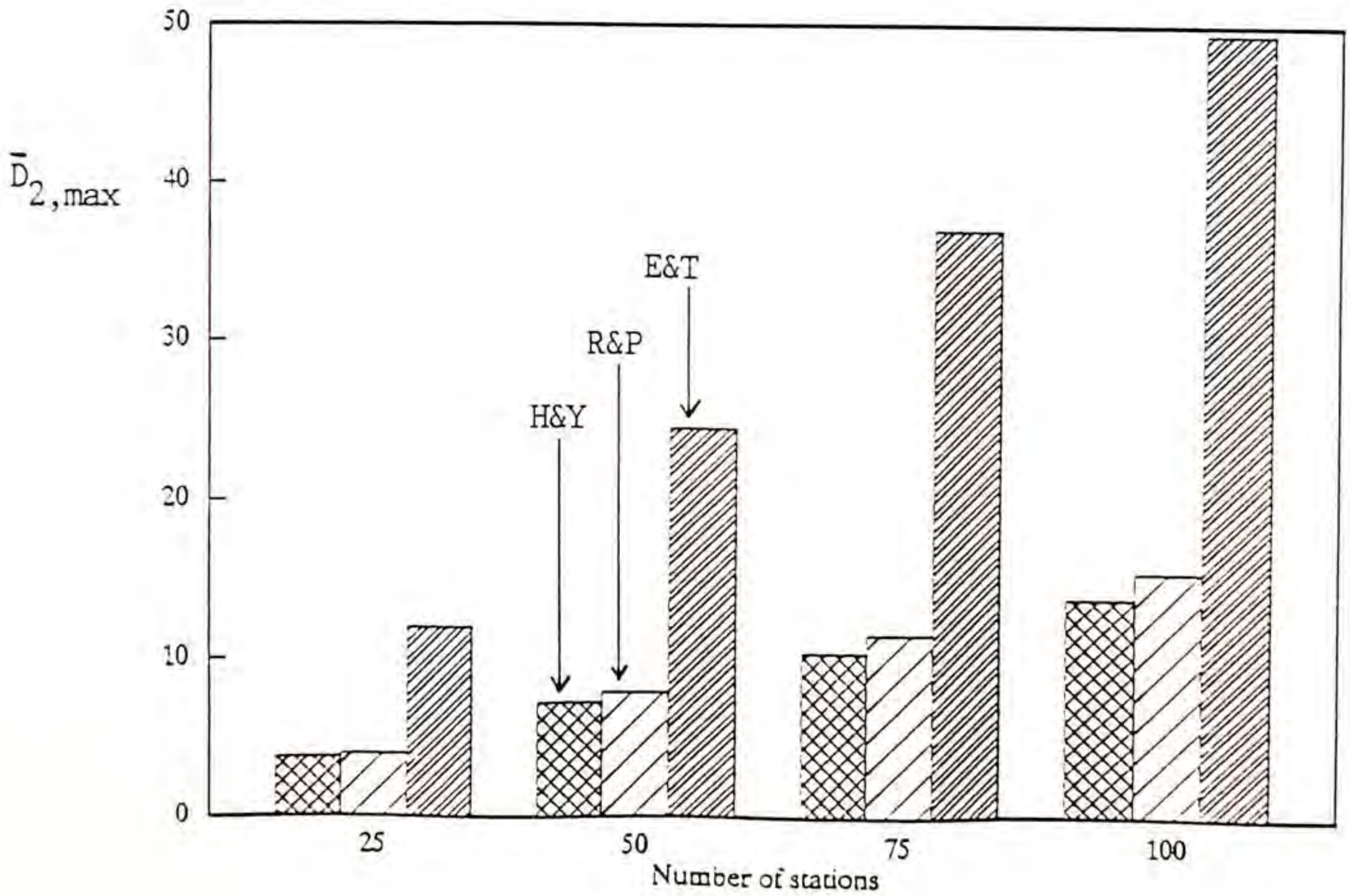


(b) under light traffic conditions

Fig.5.2 Comparison of the average scheduling delays

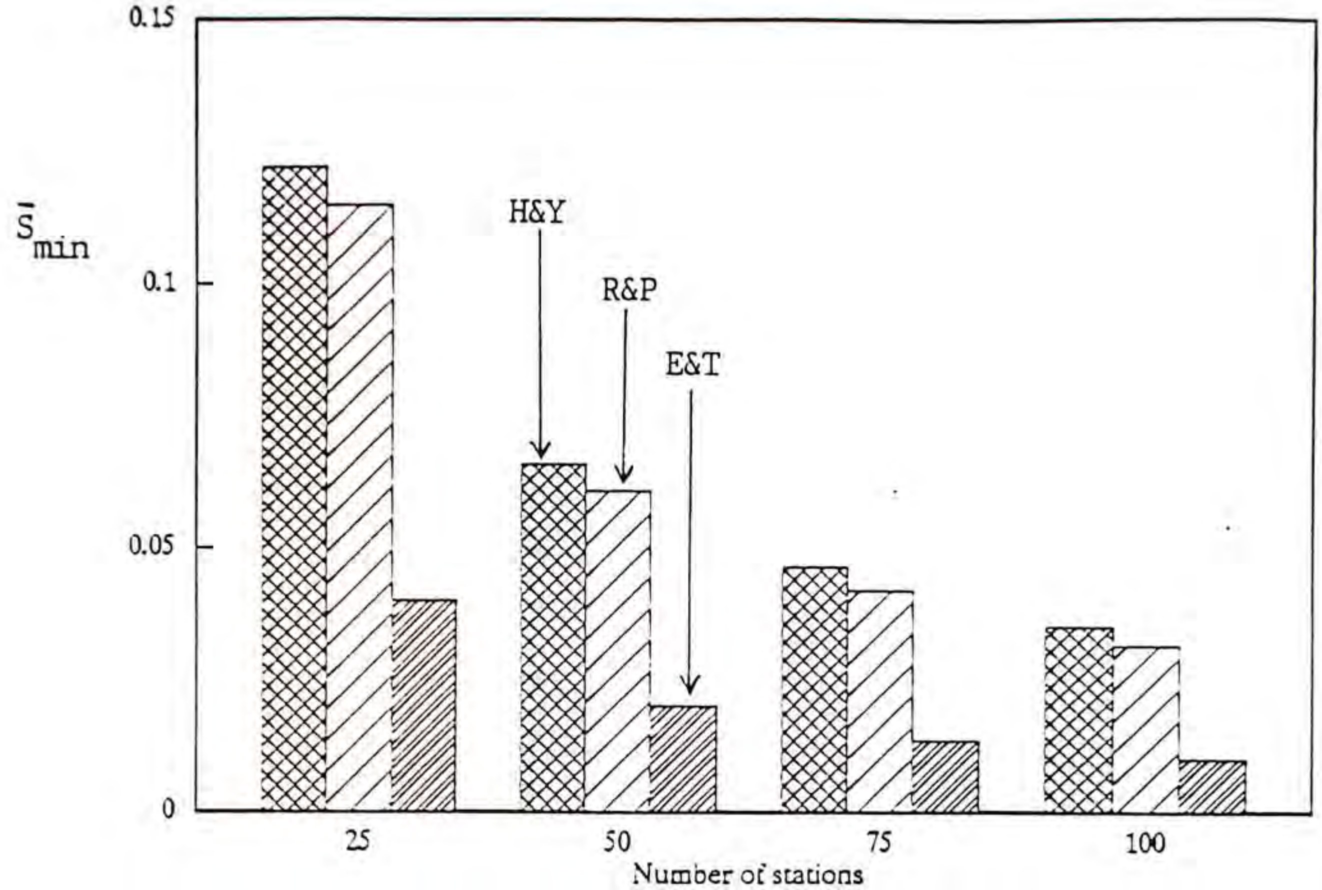


(a) under heavy traffic conditions

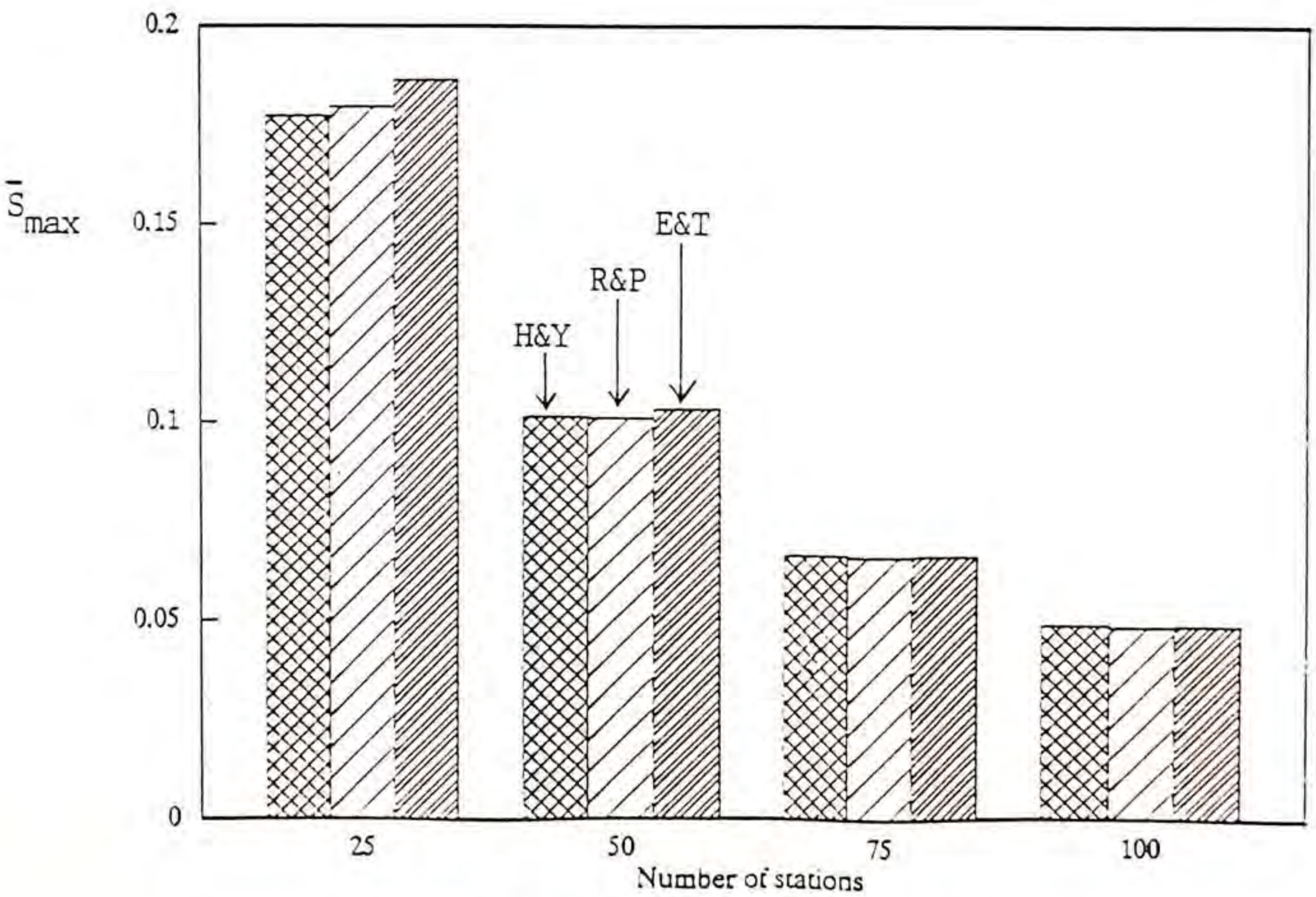


(b) under light traffic conditions

Fig.5.3 Comparison of the maximum scheduling delays



(a) the minimum transmission capacity



(b) the nomalized network capacity

Fig.5.4 Comparison of the network capacities

CHAPTER 6

STAGGERED MULTICAST PROTOCOL WITH COLLISION-FREE ACKNOWLEDGEMENT

6.2 INTRODUCTION

Broadcasting is very often used for updating distributed databases and routing tables in a communication network. The use of broadcasting in Packet Radio Networks (PRNs) is facilitated by the broadcasting nature of the medium. When the network size gets larger, a multihop network involving packet relaying is usually used for connecting all stations. After a source station has broadcast a packet, a subset of its neighbouring stations needs to rebroadcast that packet.

Very few studies of broadcasting in Spread Spectrum PRNs (SS-PRNs) is found in the literature. In spread spectrum communications, the use of spreading codes permits a receiver to extract a particular signal from many overlapping ones and adds another dimension in the design of PRNs. It is difficult to design a receiver that can simultaneously monitor all the codes. Therefore there must be rules specifying which set of codes is to be monitored and which set of codes is to be used for transmission for each station. Four types of spreading code protocols can be identified: common code protocols, receiver-based protocols, transmitter-based protocols and hybrid protocols [SOUS 88].

The use of common code protocols facilitates the transmission of broadcast packets because all stations are tuned to the common code at all times. Transmitter-based protocols are also suitable for broadcasting but the receiver must know the transmission code used in order to receive. Receiver-based protocols are not suitable for broadcasting because a separate transmission is required for each receiver. It was suggested in [PURS 87] that a fraction of the packet slots can be designated as broadcast slots using a common code while the transmission in the "non-broadcast" slots could use a receiver-based protocol. Among the hybrid protocols, the Common-Header/Transmitter-Based Protocol [SOUS 88] looks most promising for broadcasting. Here the destination and source addresses are transmitted using the common code while the data is transmitted using a transmitter-based code. With this arrangement only the header of the packet is under contention, while the remaining data portion is collision-free due to the use of a unique spreading code.

In broadcasting, as well as unicasting[1], when a packet is received without checksum error, the receiver transmits an acknowledgement (ACK) packet back to the transmitter. The transmission is considered successful only when the ACK packet is received by the transmitter within a time-out interval. In conventional PRNs, the transmission of a relaying packet can serve as an implicit acknowledgement to the previous transmitter. As noted in [PURS 87], implicit acknowledgement can also be used in transmitter-based and common code protocols for SS-PRNs with compatible transmission and routing protocols. The same cannot be true for receiver-based protocols since the relaying packets are in different codes.

There are very few studies on acknowledgement algorithms for SS-PRNs. Sasty examined the effect of acknowledgement traffic on the performance of slotted ALOHA-CDMA [SAST 84]. He assumed that the system has a central station and separate frequencies are used for inbound and outbound traffic. This allows stations to transmit and receive at the same time. Lee and Silvester studied the effect of acknowledgement on the performance of distributed single-hop SS-PRN using Slotted-ALOHA protocol [LEE 86]. The system considered uses a receiver-based spreading code for data transmission and a transmitter-based code for ACK packets. Stations cannot transmit and receive at the same time. They considered only the single destination transmission and so ACK packets are always collision-free as transmitter-based code is used.

---

1 This is commonly referred to as point-to-point transmission or single destination transmission by some authors. We choose to call it unicast so that it can easily be distinguished from broadcast and multicast.

Acknowledgement in broadcasting is quite a different problem as all neighbours of the transmitting station need to acknowledge. In conventional PRNs if the neighbours acknowledge at the same time, the ACK packets will collide. With the use of spreading code, ACK packet collision will still occur if common code or receiver-based code are used. There will be no ACK packet collision for transmitter-based code but the transmission of ACK packets by the neighbours must still be staggered in time as the station expecting ACK packets cannot monitor all the different codes of its neighbours simultaneously.

Broadcast and unicast protocols are usually designed separately. In this chapter we design the Staggered Multicast Protocol with Collision-free Acknowledgement which is suitable for unicasting, broadcasting as well as multicasting in multihop SS-PRNs. The new protocol combines the feature of transmission scheduling and Common-Header/Transmitter-Based spreading code to allow overlapping of packet transmissions. The neighbouring stations of a broadcast source are scheduled to relay broadcast packets in different "header size" minislots. Hence these packet transmissions can be staggered to give significant reduction of broadcasting delay. A multicast tree found from the routing table is used for global multicast. When a station receives a multicast packet, it is responsible for forwarding the packet only to destinations on the branch of the multicast tree spanning from that station. Thus unlike the flooding algorithm, no redundant packet is generated throughout the multicast. We also design special addressing method and packet format to allow dynamic scheduling of acknowledgement packets. Simulation result shows that the new protocol provides better throughput-delay performance than the Common-Header/Transmitter-Based Slotted ALOHA protocol despite its added capabilities of staggered relay broadcasting, collision-free acknowledgement and global packet multicast.

6.3 SYSTEM MODEL

Let the locations of the stations in a multihop packet radio network be fixed and let the transmission range be the same and fixed for all stations. Let all stations within a station's transmission range be called the neighbours of that station. Each station is assigned a spreading code and an address. The address is globally unique, but the code is unique only among other stations within two-hop distance so that beyond a certain range the codes can be reused. There are many ways to assign codes to stations. One very efficient assignment algorithm requiring only a minimum number of codes can be found in Chapter 4 and is used here for code assignment. Let all stations use the same frequency band for transmission. Stations therefore cannot transmit and receive at the same time.

The Common-Header/Transmitter-Based spreading protocol is chosen for data packets transmission. With this protocol each station is assigned with a transmission code. In addition, there is a common code which is used by all stations for addressing purpose. The packet header is transmitted using the common code while the remaining portion of the packet is transmitted using the transmitter code. A minislotted approach similar to [SOUS 88] is adopted. Let a slot be defined as the length of the packet header and let the packet length be in unit of slots. We assume that the length of an ACK packet is smaller than a slot. This assumption will be justified when we discuss the acknowledgement protocol in the next section.

Stations are allowed to transmit only at the beginning of their assigned slots. Packet transmissions are scheduled in such a way that stations transmitting in their assigned slots only will not encounter conflicts. Here we use the scheduling algorithm in Chapter 5 to do such transmission scheduling. This scheduling algorithm is found to give schedules that have the shortest cycle length, the smallest scheduling delay, the largest minimum transmission capacity and the same highest normalized network capacity when compared to two of the best scheduling algorithms in the literature.

Example: Figure 6.1 shows a 15 node network. An edge between two nodes indicates that the nodes are within the transmission range of each other. Table 6.1 shows the code assignment and transmission schedule produced by the corresponding algorithms in Chapter 4 and 5. The length of the schedule cycle is 8 slots which is the minimum since station J has 7 neighbours. The number of codes required is also the minimum. In the table a letter T in the x row and y column indicates that station y is scheduled to transmit in slot x.

6.4 PROTOCOL DESCRIPTION

Following the scheduling algorithm, stations in the network can transmit in different slots without conflict. However, such transmission cannot be received by busy stations, either busy in transmitting or in receiving another packet. To make sure the destination does receive a packet correctly, some form of acknowledgement is required. For noisy channels, acknowledgement is needed even if the transmission scheduling is collision-free.

We use two types of acknowledgement packets. A positive acknowledgement packet (ACK) is returned when the target station receives the packet correctly. A negative acknowledgement packet (NAK) is returned when the received packet contains error. When the source station receives an NAK packet, it retransmits the packet immediately. When the target station is busy, no acknowledgement packet is returned and the source station will retransmit the packet after a random delay.

Acknowledging a broadcast transmission is more complex since all the neighbours of the source station need to respond. If a subset of the neighbours fails to acknowledge, this subset will be the intended receivers when the packet is retransmitted. Transmission to a subset of neighbours is called local multicast. To accommodate local multicast the address field in the packet header needs to be expanded.

6.4.1 Packet Format

The packet format is shown in Figure 6.2. It consists of three parts, header 1 followed by header 2 and the packet body. Header 1 contains the packet I.D., the receiver code bit-map and the transmitter's code number and is transmitted using the common code. The packet I.D. is a globally unique number for identifying different packets. The receiver code bit-map indicates which neighbours are in the reception list. Since the spreading code assigned to a station is unique among the station's neighbours, this code number is in fact a local address. If a neighbour with assigned code i is the intended receiver, the i-th bit in the bit-map is set to 1. When a station with assigned code i receives a packet header with a "1" in the i-th position of the bit-map, it tunes immediately to the transmitter's code to receive the rest of the packet. The length of the bit-map is equal to the total number of codes used in the network. A 100 station network for example has a bit-map length around 25 bits. To illustrate, consider the network in Figure 6.1. If station F wants to multicast a packet to stations J, L and K, the 1st, 4th and 5th bits of the bit-map are set to 1. After noticing the 4th bit in the address bit-map is 1, station L will tune its receiver to station F's transmitting code, or code 3, as indicated in header 1.

To keep the length of header 1 (and hence the slot size) short and fixed, other address information is placed in header 2 and is transmitted using the transmitter code. For unicast packets, this information includes the destination and source addresses. For broadcast packets, a special code of all "1"s is used for identification in the destination address field. For multicast packets, a special code of all "0"s followed by a list of multicast destinations are needed. The multicast destination list has the form $D_1R_1D_2R_2D_3R_3...D_VR_V$ where V is the total number of destinations, D_i is the i-th destination address and R_i is the assigned code of the relaying station responsible for forwarding the packet to destination D_i. As variable length packets are allowed in the network, a packet length field is required. The packet body contains the data and a cyclic redundancy check (CRC) field and is transmitted using the transmitter code.

6.4.2 Global Multicast

For global multicast, fixed routing with routes defined by a routing table is assumed. All stations are also assumed to have the same routing table. The paths of a multicast packet from the source station to the final destinations form a multicast tree (found from the routing table). The source station first multicasts the packet to all its neighbours on the routing tree. When a multicast packet is received by a station, that station might have to relay the packet with an updated multicast list. The updated multicast list contains only destinations on the branch of the multicast tree spanning from that station. Broadcast packets are treated as multicast packets with the multicast tree spanning all stations in the network. Note that the receiver code bit-map in header 1 is used for local multicast while the multicast list in header 2 is for global multicast. Obviously, the last hop of all global multicast can be treated as local multicast.

6.4.3 Dynamic Scheduling of Receiver-based Acknowledgement

If the receiving stations want to send back acknowledgement packets without following the data transmission schedule, they should not use common code because in doing so these acknowledgement packets would collide with the headers of other data packets. If transmitter-based code is used for acknowledgement, the source station needs to monitor different codes from its neighbours simultaneously and is therefore also not acceptable.

The Staggered Multicast Protocol uses receiver-based code for acknowledgements so that the source station needs only to monitor its own code for detecting all acknowledgement packets from its neighbours. Since only the neighbouring stations in the reception list will send back acknowledgement packets, a local scheduling among these neighbours is sufficient to make the acknowledgement packets collision-free. With that, we can summarize the Staggered Multicast Protocol with Collision-free Acknowledgement as follows.

6.4.4 Transmission Protocol

(1) When there is a packet ready for transmission, set up the header fields as follows:

  (a) For unicast packets, fill the destination address field with the address of the final destination.

  (b) For multicast packets, fill the destination address field with all "0"s, find the relaying neighbours from the routing table and formulate the multicast destination list.

  (c) For broadcast packets, fill the destination address field with all "1"s.

(2) Set the bits corresponding to all intended receivers to 1 in the receiver code bit-map.

(3) Wait for the next scheduled slot. Transmit header 1 using the common code and switch to the local station's assigned code for the rest of the packet.

(4) Monitor the local station's assigned code in the next k slots where k is the number of intended receivers. Remark: This is for detecting returned acknowledgements.

  (a) If ACK packets are received from *all* intended receivers, end.

  (b) If NAK packets are received from some intended receivers, update the receiver code bit-map and return to step (3).

  (c) Otherwise, update the receiver code bit-map, wait for a random delay and return to step

(3).

Remark: An intended receiver will either acknowledge or not acknowledge. When an acknowledgement is sent, it could be either an ACK or an NAK packet. Thus the acknowledgement status of a set of intended receivers must be either one of the seven cases shown in Figure 6.3. These seven cases can be partitioned into 3 groups corresponding to conditions (a), (b) and (c) in step (4).

6.4.5 Reception Protocol

(1) Monitor the common code to detect packets with local destination.

(2) Identify the transmitter's code, say code X, and switch to code X to receive the remaining packet.

(3) Examine the checksum error and send either an ACK or an NAK packet using code X in the m-th slot counting from the end of the data transmission, where m is the receiver's position in the receiver code bit map.

6.4.6 Processing of Transit Packets

Packets received that are not destined for the local station need to be processed and forwarded. If the transit packet is of the unicast type, just forward it using the Transmission Protocol. If the transit packet is of the multicast type, examine the multicast list and choose all D_i's such that R_i = Y, where Y is the code of the local station. The chosen D_i's and codes of their relaying stations form the new multicast list. All broadcast packets received are converted to multicast packets with updated multicast list when forwarding is needed.

6.4.7 Illustrate Examples

To illustrate the operation of protocol consider again the network in Figure 6.1. When station I multicasts a packet to H and M, it monitors only the next 2 slots after data transmission for acknowledgement. Since only two "1"s appear in the 3rd and 6th bit position of the receiver bit-map, H and M send back acknowledgement packets in the 1st and 2nd slots respectively from the end of the received data packet. For unicast packets, k and m defined above are both 1. Therefore the acknowledgement packet is sent immediately after receiving the data packet.

Note that when the source station receives an acknowledgement packet, it knows from the slot position which neighbouring station is sending the acknowledgement packet. Even if we include the receiver address and the packet I.D. in the acknowledgement packet, it is still sufficiently small to fit into a header slot.

To illustrate the processing of global multicast consider again the network in Figure 6.1. In this network an 8 bit long bit-map is needed. Consider the case where station C multicasts a packet to stations E, H, L, N and P. Figure 6.4(a) is the routing table obtained by minimum hop routing rule and Figure 6.4(b) is the multicast tree found from the routing table. Figure 6.4(c) shows the header information of the source packet and the relaying packets. Since station F is used to forward the packet to station L, and station E is used to forward the packet to stations H, N and P, the multicast destination list formulated is E2H2L3N2P2. The destination address field is filled with all "0"s and the 2nd and 3rd bits in the receiver bit-map are set to 1 to notify stations E and F to receive the packet. When station F receives the packet, only the corresponding R_i value of destination L in the multicast list is found to match with F's assigned code (i.e. code 3). Station F therefore converts the packet to a unicast type and forward it to station L. When station E receives the packet from station C, the new multicast list formulated is H5N1P1. Thus the 1st and 5th bits in the receiver bit-map are set to 1 and the packet is forwarded to stations D and J. Finally station D forwards the packet to station H and station J forwards the packet to stations N and P and the multicast process is completed.

6.5 STAGGERED RELAY BROADCASTING

After a station has broadcast a packet, a subset of its neighbours needs to rebroadcast that packet. In conventional PRNs this subset of neighbours will have to randomize their rebroadcasting time to minimize collision. If conflict-free scheduling is used, these neighbours will rebroadcast one after the other in different "packet size" slots and so the broadcasting delay, i.e. the time required for the broadcast packet to be received by all stations in the network, will be very long. The Staggered Multicast Protocol allows neighbouring stations to start transmission in different "header size" slots and thus significantly reduces the broadcasting delay.

We use the network in Figure 6.1 again to illustrate the staggering operation. In the following examples we assume that the broadcasting of a single packet from a source station to all other stations is the only activity in the network. In addition, an error-free channel is assumed and a schedule cycle of 8 slots is used. Let $S \rightarrow (D_1, D_2, ...)$ denotes the broadcasting of a packet by source station S to stations $D_1, D_2, ...$ where these neighbours are receiving the first copy of the packet. The special case $S \rightarrow ()$ occurs when all target stations are either busy or have already received the packet before and they therefore do not tune to the transmitter code of S. For simplicity the acknowledgement packet is not shown in the examples.

Consider the case of broadcasting a packet from station E to all other stations. Figure 6.5(a) is the broadcast tree found from the routing table in Figure 6.4(a). Figure 6.5(b) shows the sequence of the staggered relay transmissions using the Staggered Multicast Protocol. A packet length of 10 slots is assumed. From Table 6.1, E's transmission slot is at slot 2. Starting at slot 2, E's transmission will end at slot 3 of the next schedule cycle. After receiving E's transmission, station I forwards the packet to M in slot 4 (from Table 6.1). Since D, H and J (neighbours of I) are not in the reception list, they will not switch to I's code after checking the packet header. In slot 5, D forwards the packet to H. Note that when H (which is a neighbour of D and I) receives D's transmission (i.e. monitoring D's code) it is not affected by I's transmission in I's code. Subsequently B forwards the packet to A in slot 6 and C forwards the packet to G in slot 7. Then J multicasts the packet to N and P in slot 1 of the next cycle. Finally F multicasts the packet to K and L in slot 3 and the broadcast is completed using a total time of 27 slots.

Figure 6.6 shows the sequence of broadcasting from E to all other stations using conventional radio signal (without spreading codes). Here conflict-free scheduling is chosen for packet transmission and the same transmission schedule in Table 6.1 is used. Note that the slot size is now equal to the packet length. The broadcast starts in "packet size" slot 2 and covers the whole network after stations B and M rebroadcast in "packet size" slot 6. Hence the broadcasting delay required is 5 x 10 (the packet length) = 50 "header size" slots, which is almost twice as much as what is required by staggered relay broadcasting.

Figure 6.7 shows the sequence of staggered relay broadcasting from E again but packet lengths of 6 and 100 slots are now assumed. While the broadcasting delays in a conventional PRN with conflict-free scheduling are 30 and 500 slots for the two cases, the staggered relay broadcasting needs only 18 and 207 slots to complete the broadcast. It can be seen that more reduction of broadcasting delay is obtained with longer packet size.

6.6 SIMULATION RESULTS

The performance of the Staggered Multicast Protocol (SMP) is studied by simulation on networks in a square region of dimension 20km x 20km and a transmission range of 5 km. Random station distribution and lattice networks are considered. The packet generation rates are the same for all stations and the packet destinations are equally probable for all stations. Poisson arrival of packets to all stations is assumed and minimum hop routing is used.

The average end-to-end delay as a function of network throughput for networks with random station distribution is plotted in Figure 6.8. Also shown for comparison is the Slotted ALOHA protocol using Common-Header/Transmitter-Based spreading code (CT-ALOHA). The number of stations in the network varies from 25 to 100 and the packet length is chosen as 100 slots. It is seen that the Staggered Multicast Protocol always has a better throughput-delay performance than the CT-ALOHA protocol. Figure 6.9 shows the throughput-delay performance for lattice networks. The maximum network throughout attained by SMP is found to be 10 to 15% higher than that of CT-ALOHA in both random station distribution and lattice networks. We observe that the improvement is greater for denser networks.

We then fix the number of stations in the lattice network to 100 and run the simulation for different packet lengths. The cycle length of SMP under this case is 26 slots. Figure 6.10 shows the throughput-delay performance when the packet length is 13 and 25 slots. When the packet length is only half of the cycle length SMP has very little improvement over CT-ALOHA. This is obvious because the overhead introduced by the scheduling delay is relatively higher. When the packet length is approximately the same as the cycle length the maximum network throughput attained by SMP is 17% higher than that of CT-ALOHA. It is also reasonable to observe that there is a greater improvement for networks with longer packet length.

6.7 CHAPTER SUMMARY

Broadcast and unicast protocols are usually designed separately in multiple PRNs where packet relaying is required to send a packet from a source to the stations further apart. By properly schedule transmission times, a lot of transmission conflicts can be avoided especially when packet broadcasting and explicit acknowledgements are required. The use of spread spectrum adds another dimension to the design of such system as now limited interference is allowed. We tie all these together and have designed the Staggered Multicast Protocol with Collision-free Acknowledgement which is suitable for unicasting, broadcasting as well as multicasting. The Common-Header/Transmitter-Based spreading protocol is chosen for data packets transmission and so overlapped transmissions of packet bodies are allowed. This staggering of transmission can significantly reduce broadcasting delay.

We have also designed special addressing method and packet format to achieve collision-free acknowledgement and multicasting capability. The receiver-based spreading code is used for acknowledgement packets and a dynamic acknowledgement scheduling of the neighbouring stations has been designed. Simulation result shows that the new protocol provides better throughput-delay performance than the Common-Header/Transmitter-Based Slotted ALOHA protocol.

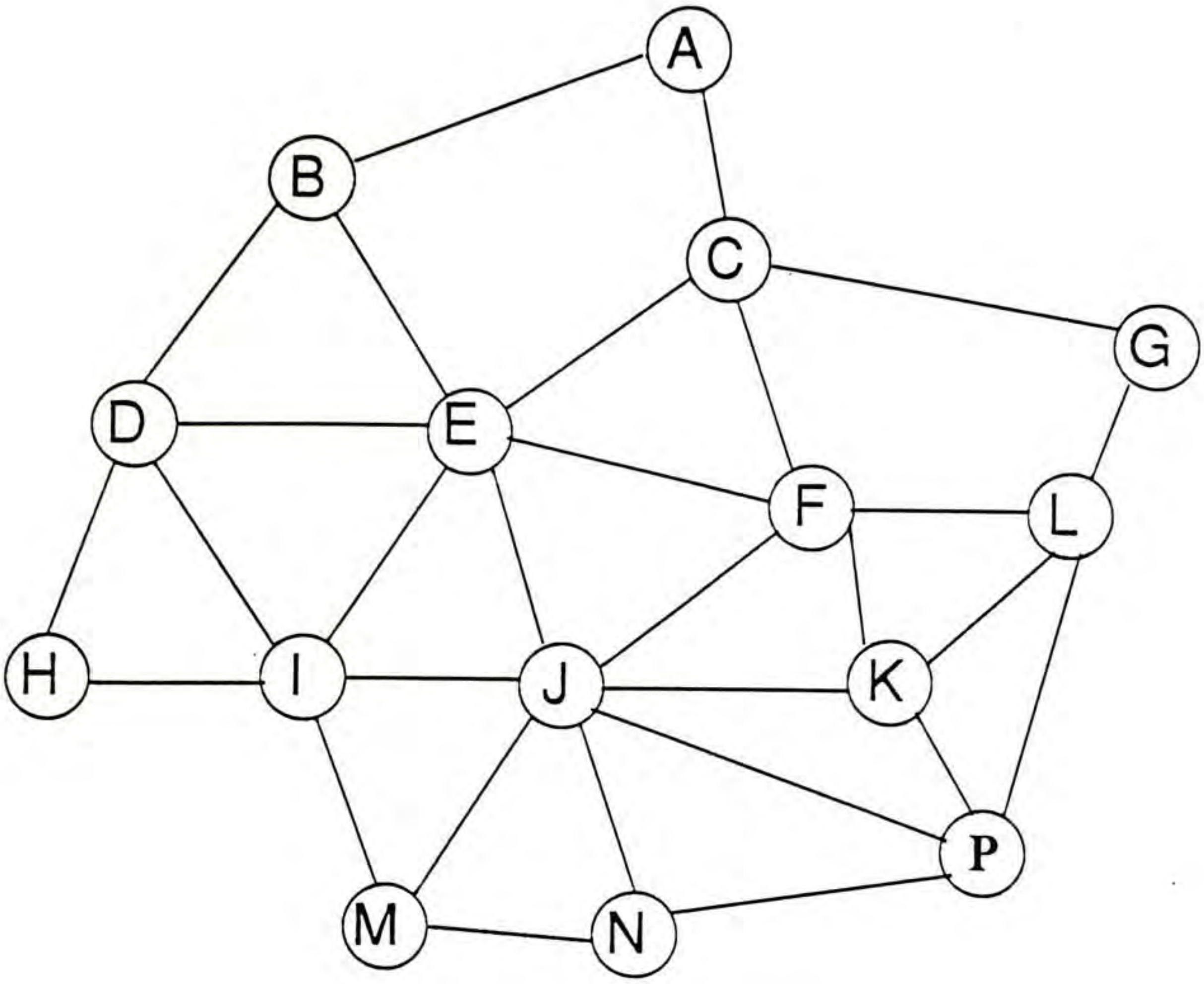


Fig.6.1 A sample network

Station (Code) / Slot Number	A (1)	B (6)	C (8)	D (5)	E (2)	F (3)	G (6)	H (3)	I (4)	J (1)	K (5)	L (4)	M (6)	N (7)	P (8)
1	T									T					
2					T										
3						T		T							
4	T								T			T			
5				T							T				
6		T					T						T		
7			T					T						T	
8			T					T							T

Table 6.1 The code assignment and transmission schedule

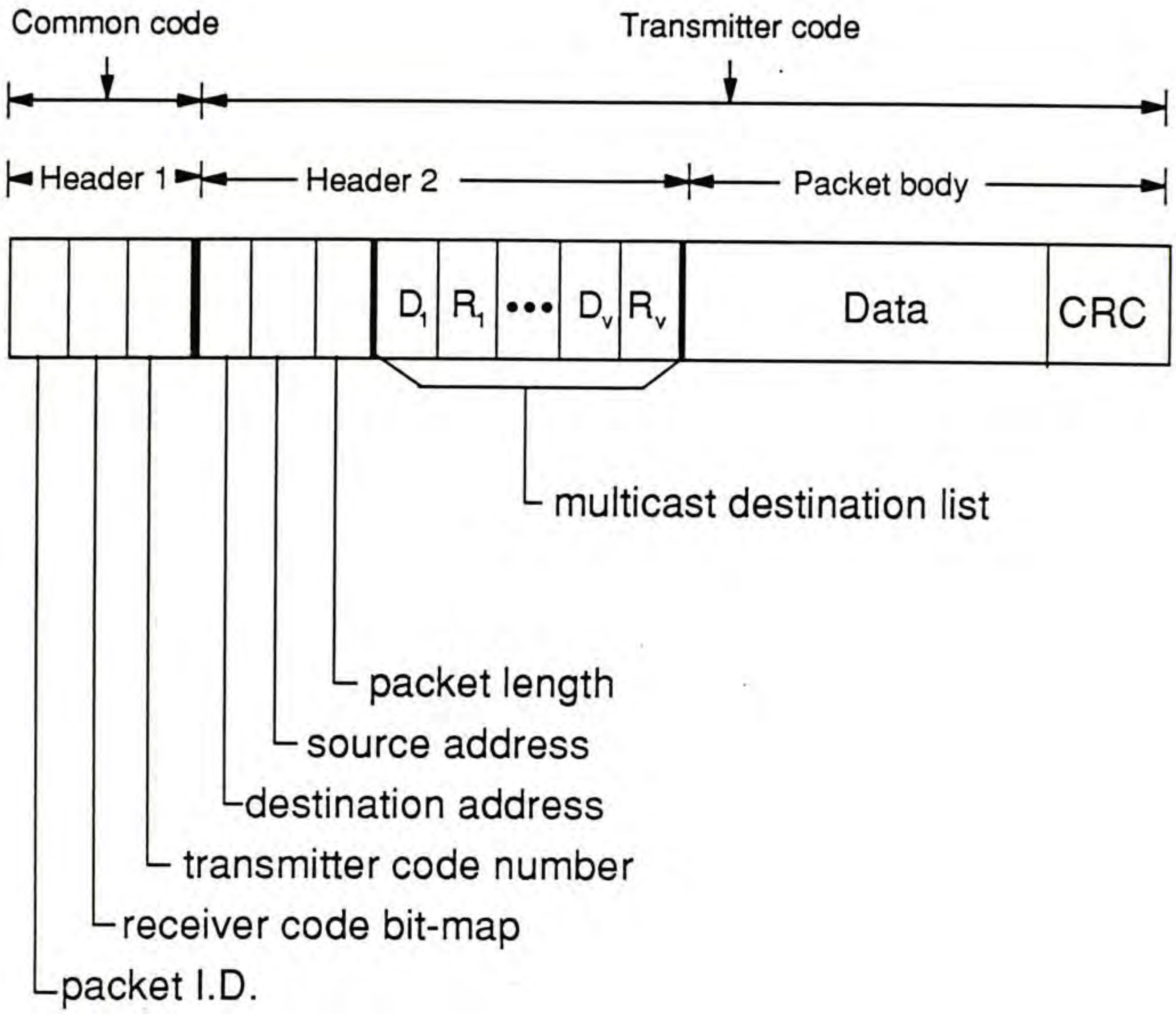
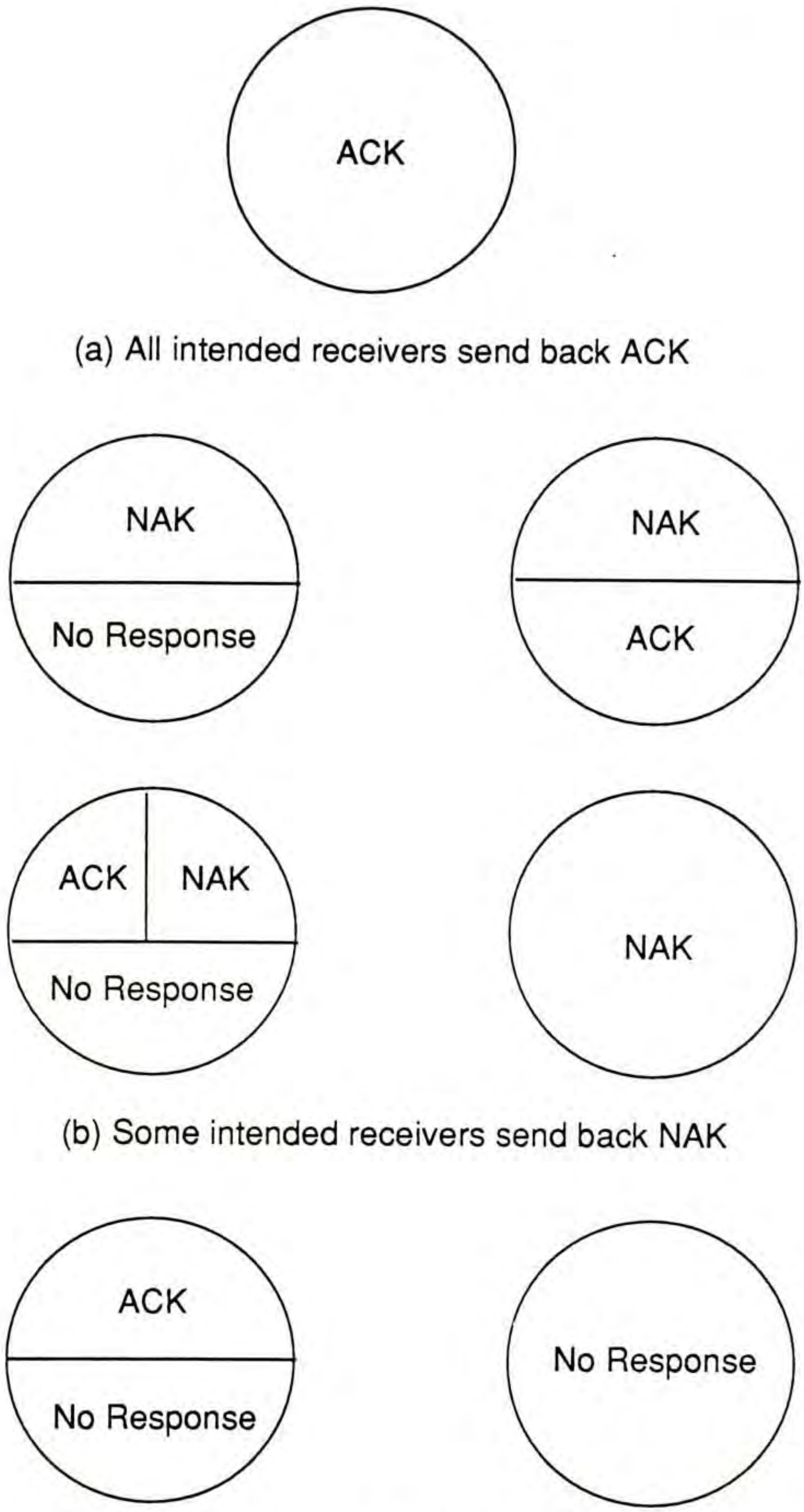


Fig.6.2 The packet format



(a) All intended receivers send back ACK

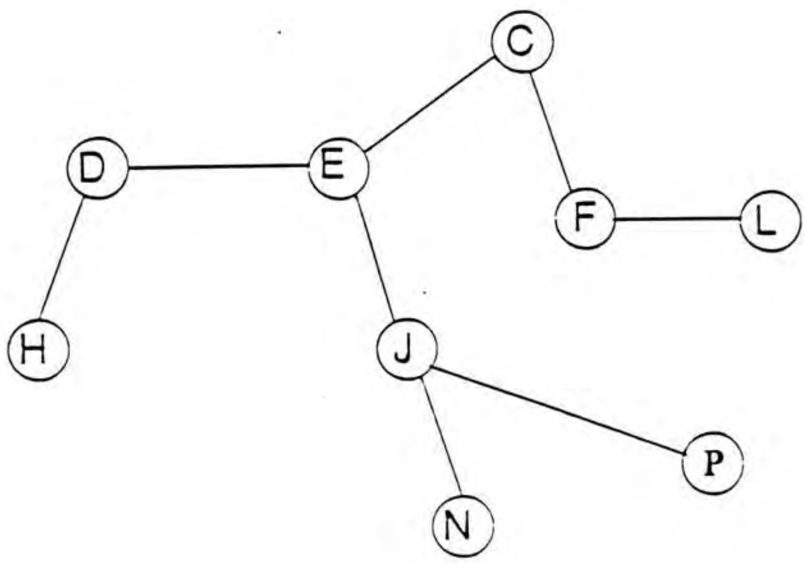
(b) Some intended receivers send back NAK

(c) Not all ACKs are sent and no NAK is sent

Fig.6.3 Cases of acknowledgement status.

Source ID \ Destination ID	A	B	C	D	E	F	G	H	I	J	K	L	M	N	P
A	-	B	C	B	B	C	C	B	B	B	C	C	B	B	B
B	A	-	A	D	E	E	A	D	D	E	E	E	D	E	E
C	A	A	-	E	E	F	G	E	E	E	F	F	E	E	E
D	B	B	E	-	E	E	E	H	I	E	E	E	I	E	E
E	B	B	C	D	-	F	C	D	I	J	F	F	I	J	J
F	C	E	C	E	E	-	C	E	E	J	K	L	J	J	J
G	C	C	C	C	C	C	-	C	C	C	L	L	C	L	L
H	D	D	D	D	D	D	D	-	I	I	I	D	I	I	I
I	D	D	E	D	E	E	E	H	-	J	J	E	M	J	J
J	E	E	E	E	E	F	E	I	I	-	K	F	M	N	P
K	F	F	F	F	F	F	L	J	J	J	-	L	J	J	P
L	F	F	F	F	F	F	G	F	F	F	K	-	F	P	P
M	I	I	I	I	I	J	I	I	I	J	J	J	-	N	J
N	J	J	J	J	J	J	P	J	J	J	J	P	M	-	P
P	J	J	J	J	J	J	L	J	J	J	K	L	J	N	-

(a) The routing table

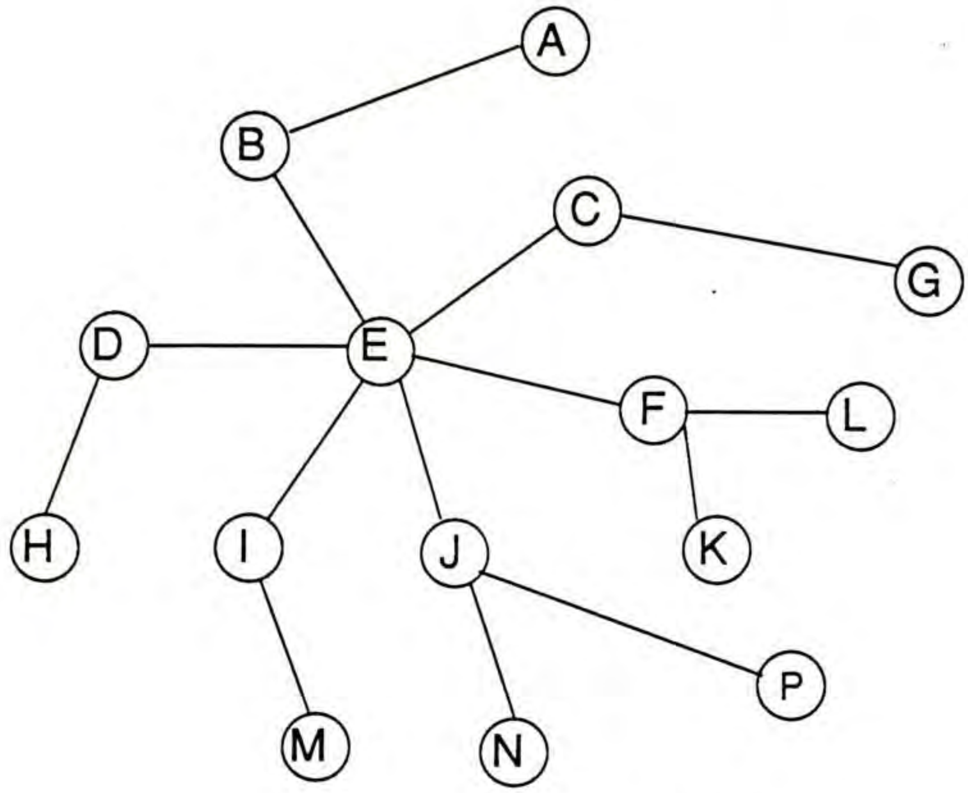


(b) The multicast tree

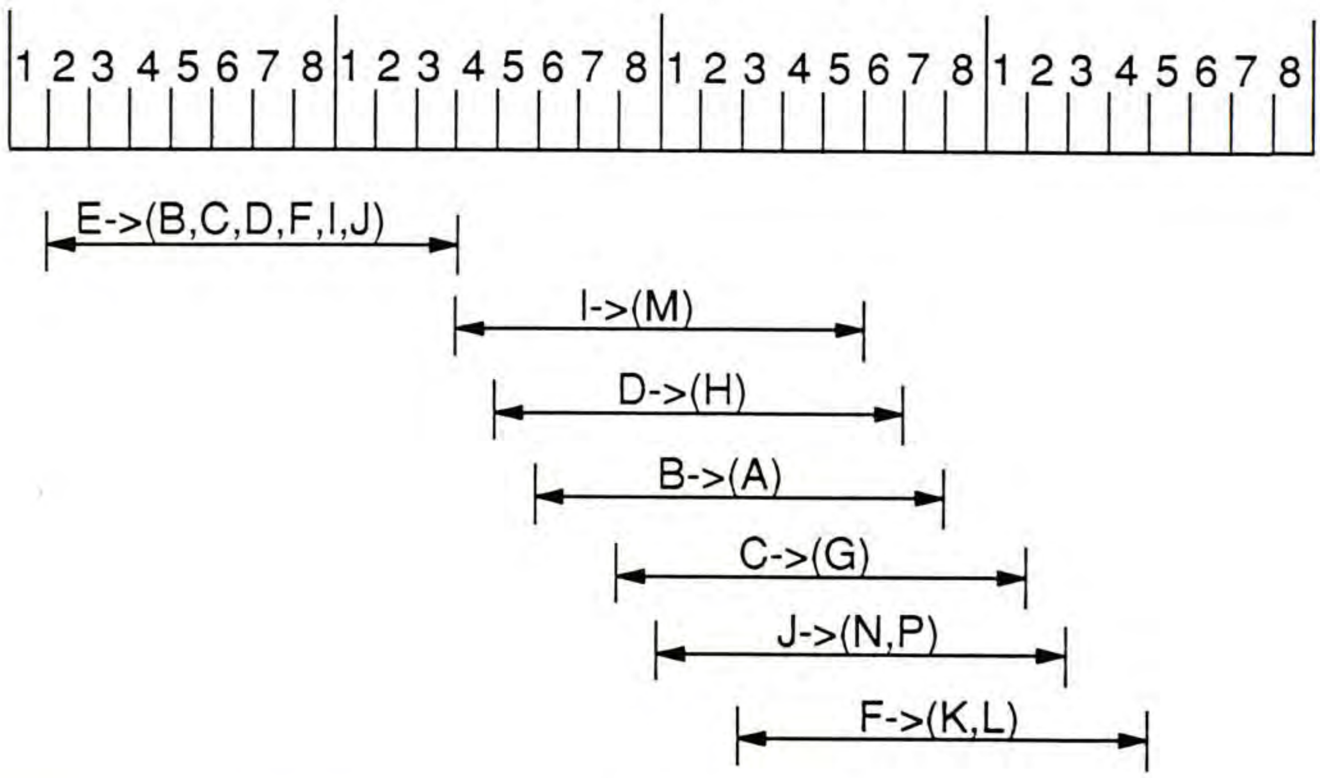
station(code)	receiver bit-map	destination address	multicast destination list
C(8)	01100000	00000000	E2 H2 L3 N2 P2
F(3)	00010000	L	
E(2)	10001000	00000000	H5 N1 P1
D(5)	00100000	H	
J(1)	00000011	00000000	N7 P8

(c) The header information

Fig.6.4 Station C multicasts a packet to E, H, L, N and P.



(a) The broadcast tree of E



(b) The staggered transmission sequence

Fig.6.5 Staggered relay broadcasting, packet length=10 slots

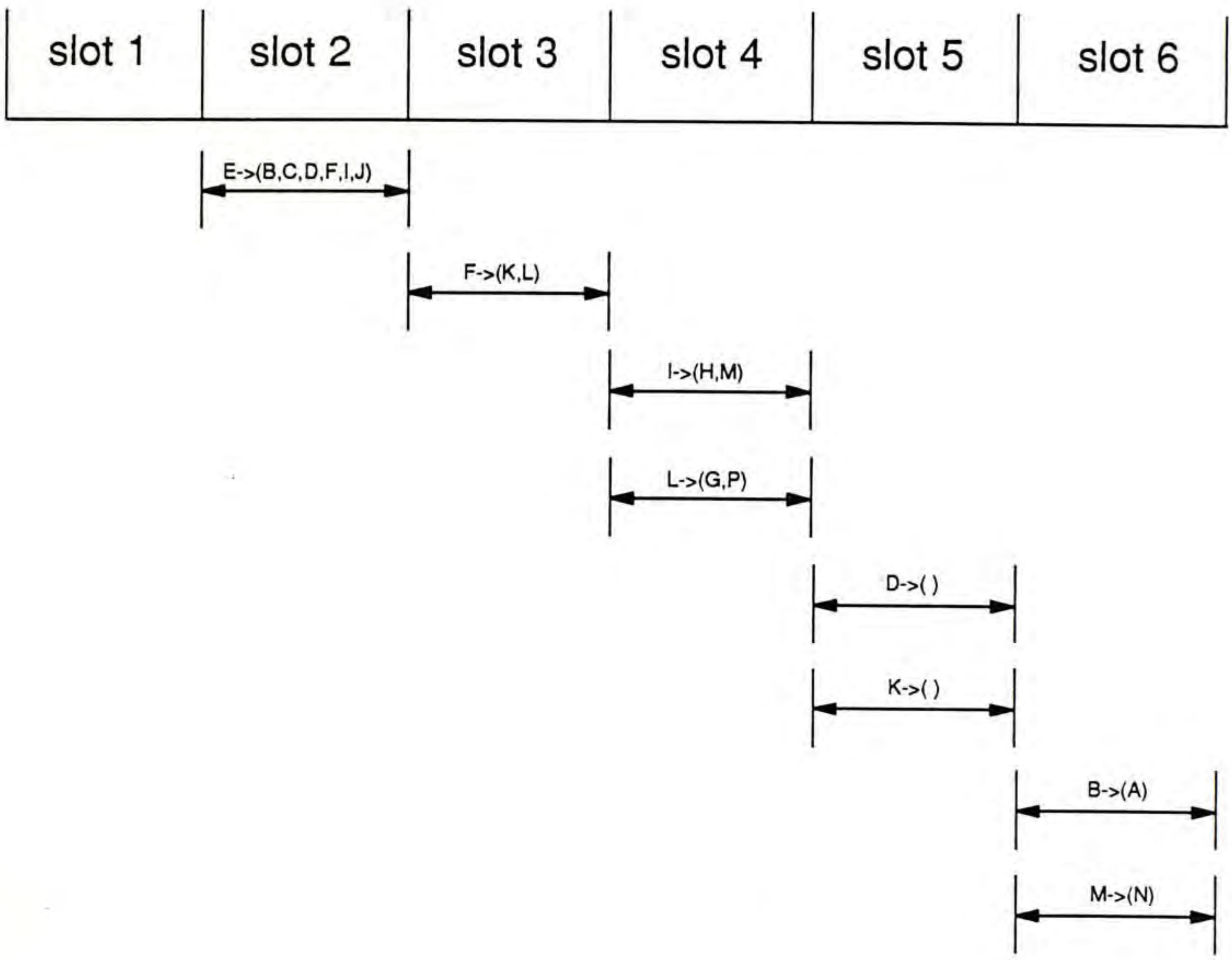
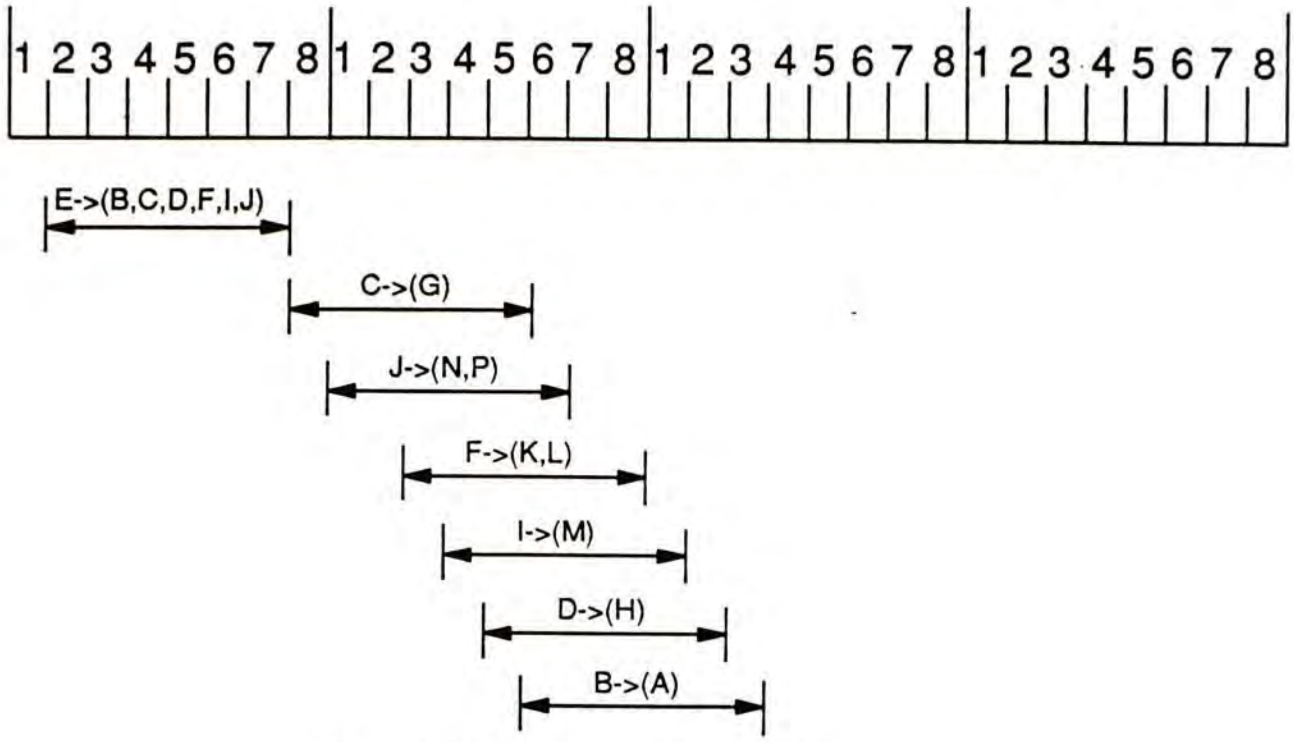
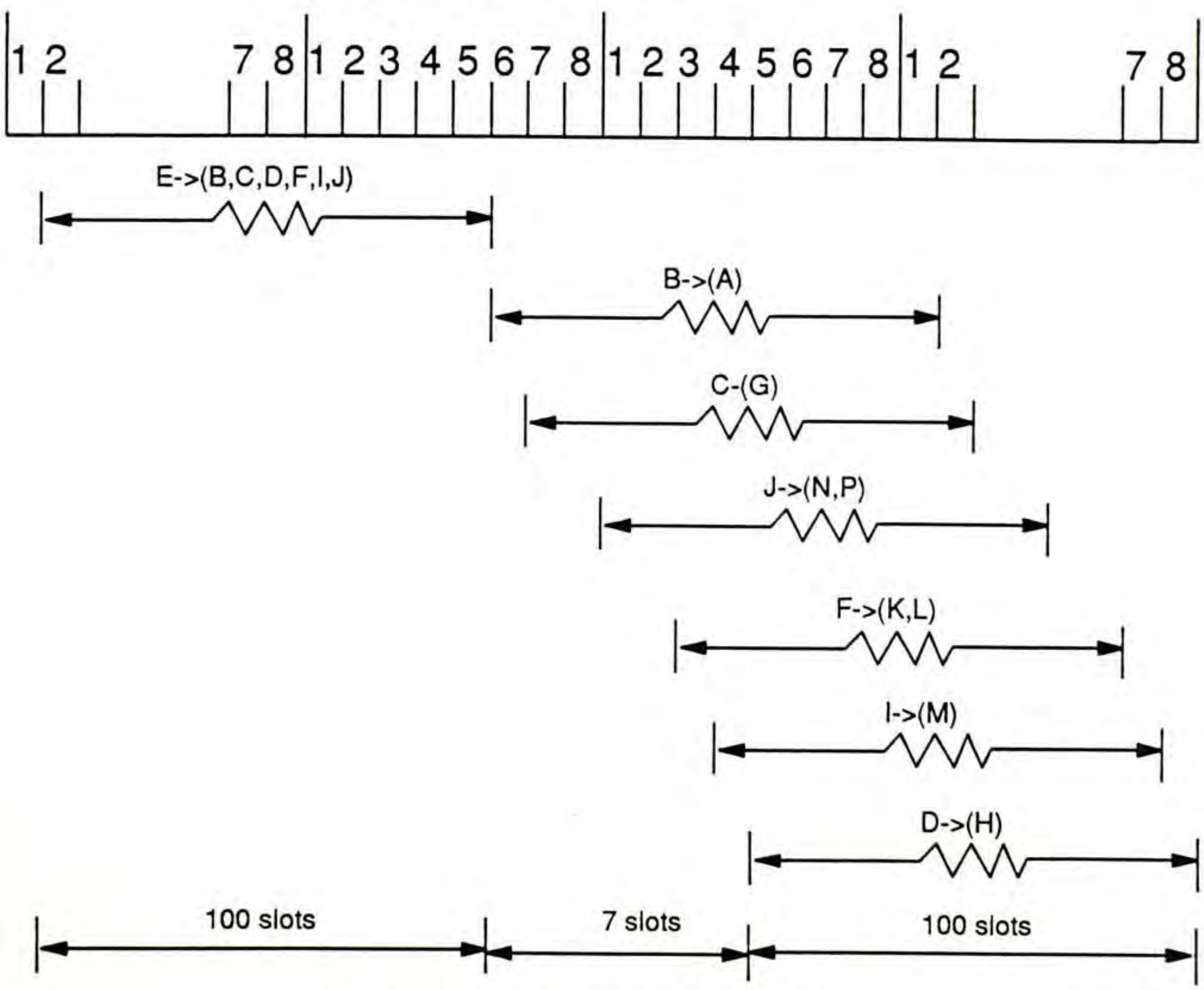


Fig.6.6 Conflict-free relay broadcasting



(a) packet length=6 slots



(b) packet length=100 slots

Fig.6.7 Staggered relay broadcasting

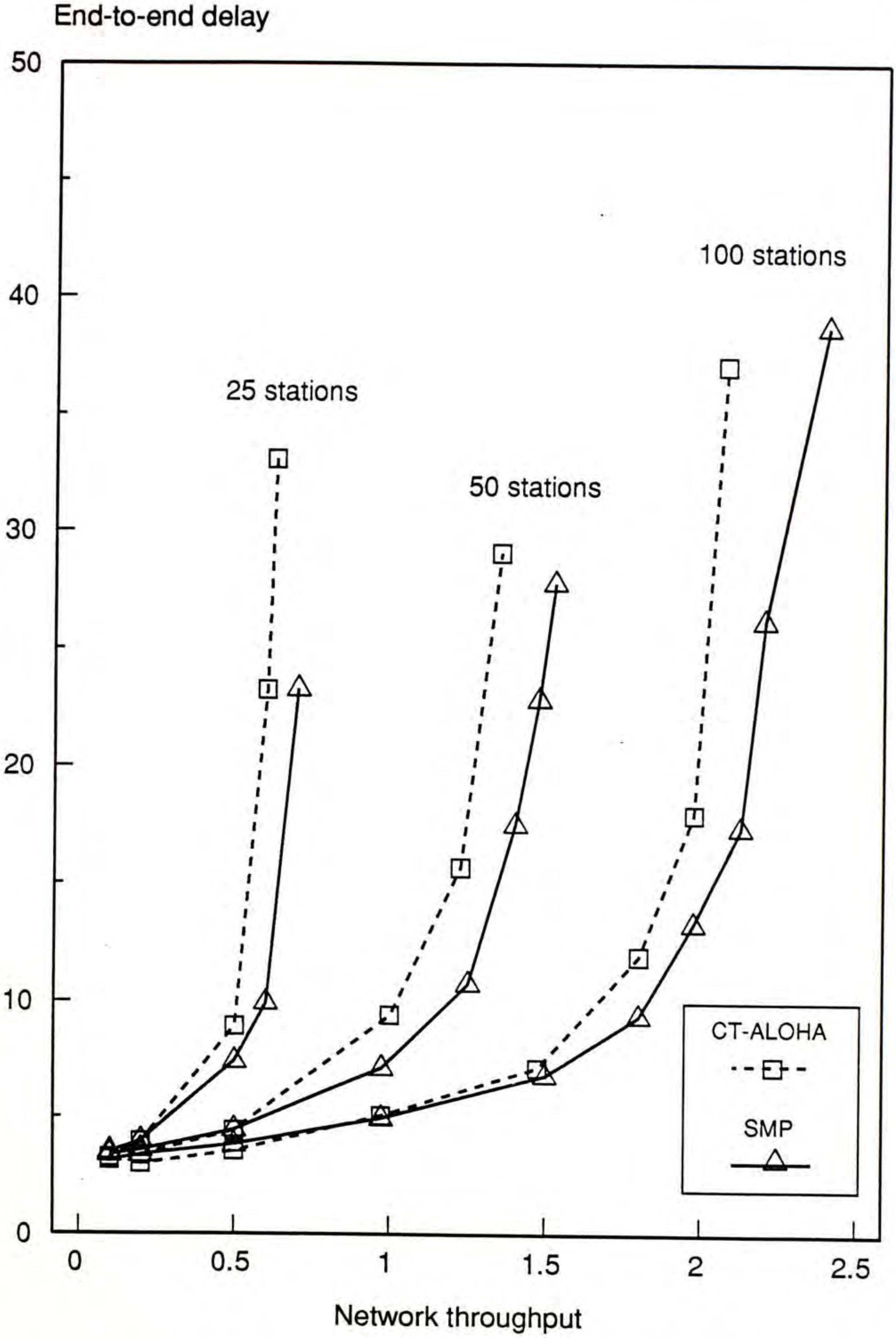


Fig.6.8 Performance comparison on random station distribution network

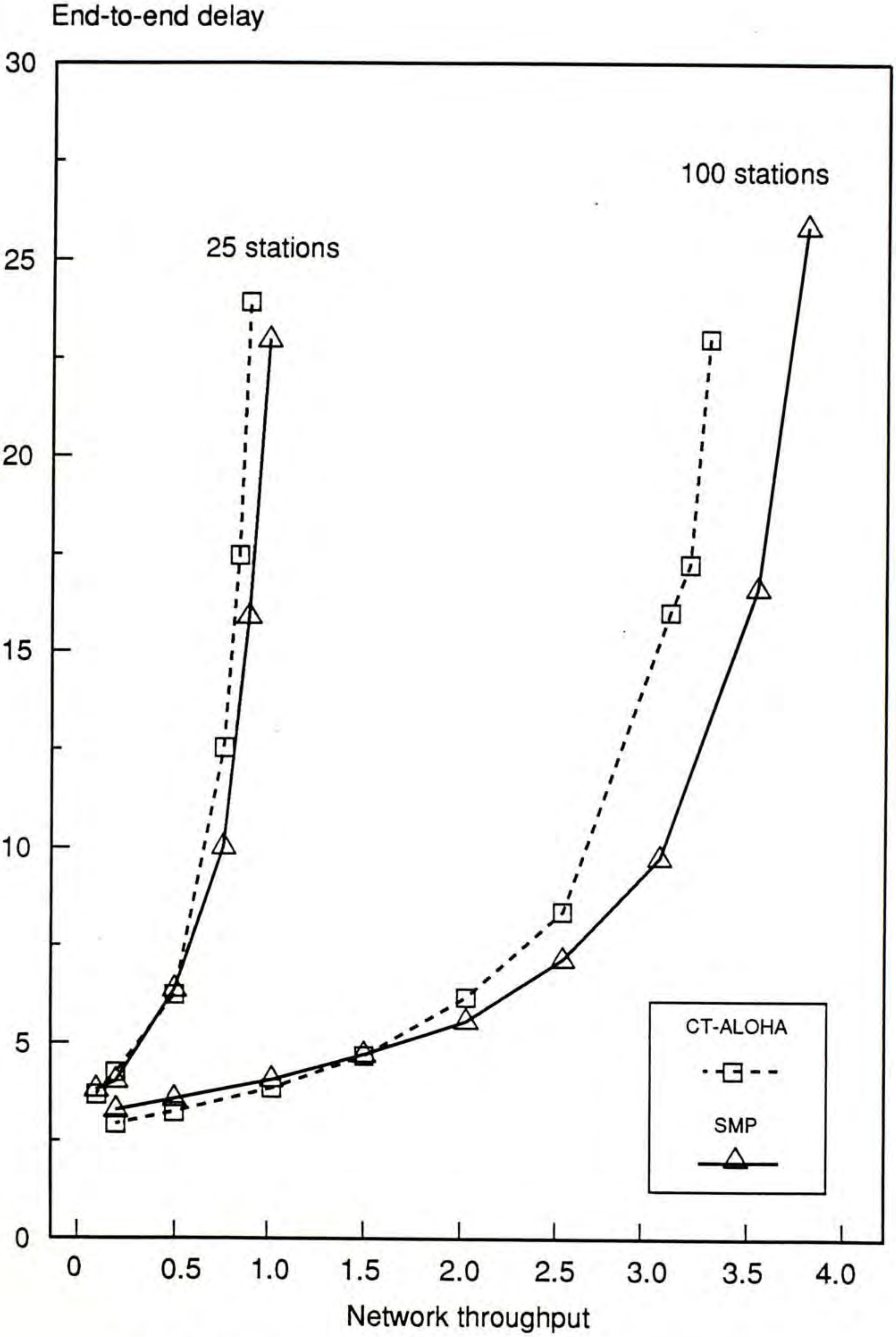


Fig.6.9 Performance comparison on lattice network.

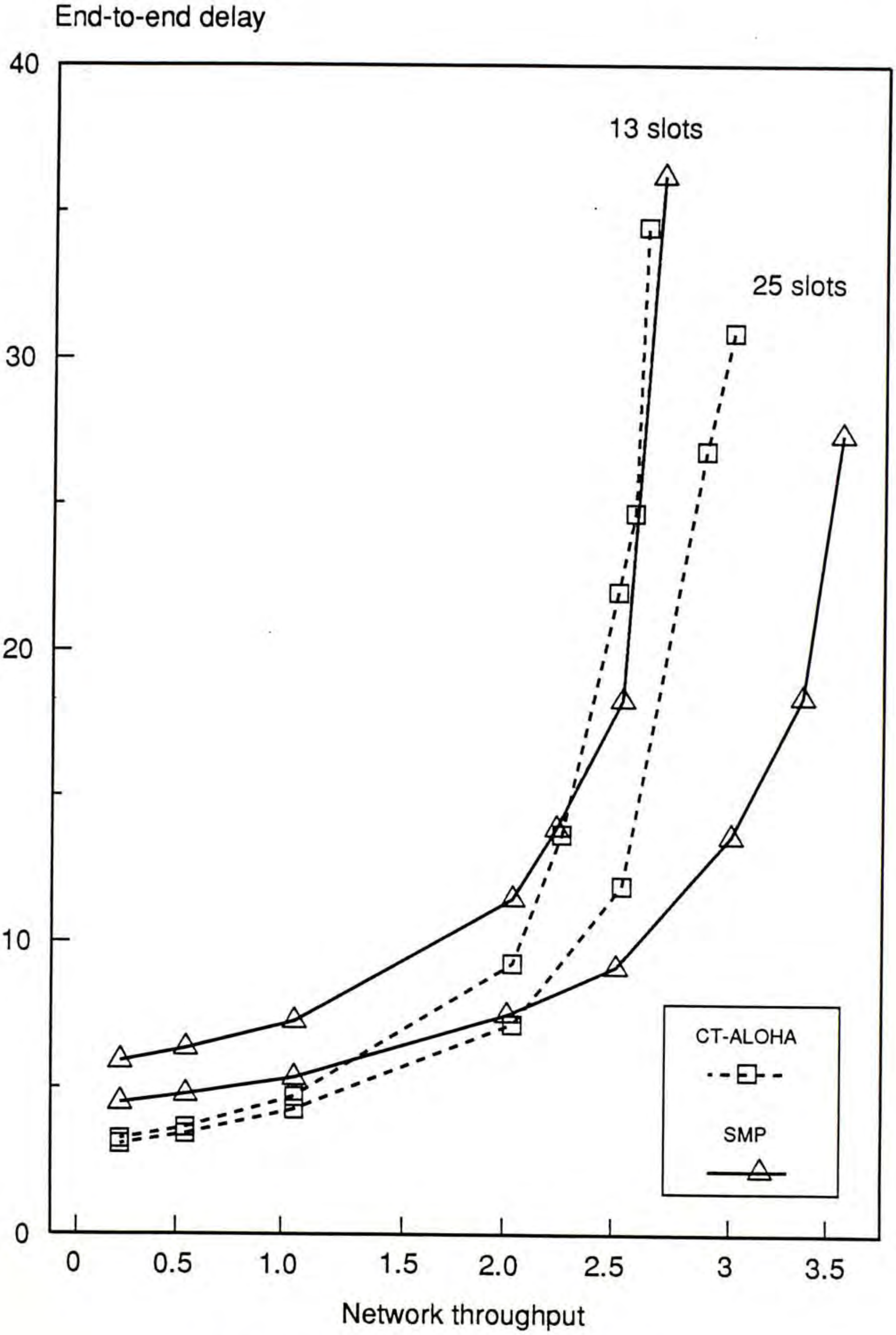


Fig.6.10 Effect of varying packet length

CHAPTER 7

CONCLUSION

7.1 SUMMARY

In this thesis, we have designed and evaluated several communication protocols and design algorithms in multihop Packet Radio Networks. We have investigated the use of directional antennas in PRNs so as to enhance the spatial reuse advantage. New unicast and multicast protocols have also been designed for Spread Spectrum Packet Radio Networks (SS/PRNs). We have also considered how to achieve fair and efficient transmission scheduling in PRNs.

In Chapter 1, we gave an overview of Packet Radio Networks. Three most commonly used channel access protocols were discussed and compared. We highlighted the design issues related to spatial reuse and spread spectrum. This overview provided the background for discussion in the following chapters.

We proposed in Chapter 2 a design methodology as well as two efficient channel access protocols for multihop PRNs with multiple directional antennas stations. The Simple Tone Sense (STS) protocol can minimize transmission interference by using a group of tones to identify the active neighbours. A variation of this protocol namely the Variable Power Tone Sense (VPTS) protocol can further reduce transmission interference by using minimum required transmission power. Algorithms for assigning tones and for determining the orientation and broadcasting angles of the directional antennas were designed. Simulation result shows that the STS and VPTS protocols performs particularly well when the traffic is heavy.

In Chapter 3, we designed a code assignment algorithm for SS/PRNs. This algorithm can reduce the number of spreading codes required to 20% - 35 % of the total number of stations in the network. We also designed the Coded Tone Sense (CTS) protocol which can further reduce the number of codes required. Simulation result shows that this protocol is particularly attractive for densely populated networks. For these networks only a few codes is sufficient to drive the throughput-delay performance very close to the case where each station has a unique code.

It is important to find an efficient algorithm for assigning as few codes to the SS/PRN stations as possible. In Chapter 4, we transformed the code assignment problem to the familiar graph coloring problem and designed a very efficient code assignment algorithm. A very tight lower bound on the number of codes needed was derived. The performance of this algorithm was assessed through extensive case studies by making comparisons to the bound as well as to one of the best heuristics for graph coloring.

In Chapter 5, we designed a very efficient scheduling algorithm for PRNs and derived (1) the schedule cycle length, (2) the scheduling delays, (3) the minimum transmission capacity and (4) the normalized network capacity as performance measures. Extensive case studies show that the new algorithm always gives the shortest schedule cycle, the smallest mean scheduling delay and the largest minimum transmission capacity when compared to two of the best scheduling algorithms in the literature. This algorithm also gives the smallest difference between the normalized network capacity and minimum transmission capacity. The normalized network capacities obtained by the three algorithms are almost identical indicating that all three algorithms are efficient and the new algorithm is more fair in capacity distribution.

Broadcast and unicast protocols are usually designed separately in multiple PRNs. In Chapter 6, we designed the Staggered Multicast Protocol which is suitable for unicasting, broadcasting as well as multicasting in SS/PRNs. The Common-Header/Transmitter-Based spreading protocol was chosen for data packets transmission and so overlapped transmissions of packet bodies are allowed. This staggering of transmission significantly reduces broadcasting delay. We also designed special addressing method and packet format to achieve collision-free acknowledgement and multicasting capability. The receiver-based spreading code was chosen for acknowledgement packets and a dynamic acknowledgement scheduling of the neighbouring stations was designed. Simulation result shows that the new protocol provides better throughput-delay performance than the Common-Header/Transmitter-Based Slotted ALOHA protocol despite its added capabilities of staggered relay broadcasting, collision-free acknowledgement and global packet multicast.

7.2 TOPICS FOR FUTURE RESEARCH

Packet Radio Networks are useful for communication in regions where wire connection between users is not practical or expensive. The work in this thesis explores several area to improve the performance of networks with fixed stations. Since wireless networks are particular suitable for communication among mobile users, it is worthwhile to investigate how to apply the idea introduced in this thesis to mobile environment. We highlight some possible research areas which are related to the work in this thesis.

(1) The use of directional antennas in a mobile network needs a different protocol from a fixed station environment. We need to identify the location of the mobile station for chosing the appropriate directional antenna. The newly designed protocol also needs to solve the problem where a receiving station is moving away from the transmission region of the source station.

(2) Almost all the existing routing algorithms are designed for use with omnidirectional antennas. New routing algorithms should be designed to make efficient use of the multiple directional antennas equipped in a station so as to resolve the congestion when the traffic is heavy.

(3) In the Coded Tone Sense protocol, we choose, for simplicity, to allocate codes so that the number of stations sharing a code is as even as possible. Generally the criteria of sharing codes depends on the station distribution and the traffic on the network. A more systematic and efficient code sharing rule is needed for stations with non-uniform traffic.

(4) To investigate the possible way of reducing the number of spreading codes required in a mobile SS/PRN.

(5) One of the original benefits of spread spectrum signaling is the secure communication between two stations. When some stations are allowed to share the same spreading code, a security problem then occurs. We need to design a new protocol which can allow sharing of codes and at the same time provide secure communications.

REFERENCE

[ABRA 70] N. Abramson, "The ALOHA system -- Another alternative for computer communications," Proc. Fall Joint Comput. Conf., 1970, vol.37, pp. 281-285.

[APPE 76] K. Appel and W. Haken, "Every planar map is 4-colorable," Bull. Am. Math. Soc. 82 (1976), pp.711-712.

[BRAZ 85] J.M. Brazio and F.A. Tobagi, "Throughput analysis of spread spectrum multihop packet radio networks," Proc. INFOCOM'85, Washington, March 1985.

[BREL 79] D. Brelaz, "New methods to color the vertices of a graph," Comm. of the ACM, Vol.22, No.4, pp.251-256, April, 1979.

[CHAN 84] J.F. Chang and C.J. Chang, "Optimal design parameters in a multihop packet radio network using random access techniques," Proc. of GLOBECOM'84, pp.493-497, Nov. 1984.

[CHEN 85] M.S. Chen and R. Boorstyn, "Throughput analysis of code division multiple access (CDMA) multihop packet radio networks in the presence of noise," Proc. INFOCOM'85, Washington, March 1985.

[CHLA 85] I. Chlamtac and S. Kutten, "On broadcasting in radio networks -- problem analysis and protocol design," IEEE Trans. Commun. vol.COM-33, No.12, pp.1240-1246, Dec. 1985.

[CHLA 87] I. Chlamtac and S. Kutten, "Tree-based broadcasting in multihop radio networks," IEEE Trans. Comput. vol.C-36, No.10, pp.1209-1223, Oct. 1987.

[EPHR 90] A. Ephremides and T. Truong, "Scheduling broadcasts in multihop radio networks," IEEE Trans. Commun. vol.COM-38, No.4, pp.456-460, April 1990.

[GERA 88] D.P. Gerakoulis, T.N. Saadawi and D.L. Schilling, "A channel access protocol for embedding CSMA on spread spectrum packet radio networks," Proc. ICC'88, Philadelphia, June 1988.

[HOU 86] T.C. Hou and V.O.K. Li, "Transmission range control in multihop packet radio networks," IEEE Trans. Commun. vol.COM-34, No.1, pp.38-44, Jan. 1986.

[HUNG 86a]K.W. Hung and T.S. Yum, "Slotted ALOHA for multihop packet radio networks with multiple directional antennas stations," Electronics Letters, vol.22, No.19, pp.1027-1029, Sept. 1986.

[HUNG 86b]K.W. Hung and T.S. Yum, "The multi-tone multiaccess protocol with collision detection for multihop packet radio networks with multiple directional antennas stations," Proc. of ICC'86, Toronto, Canada, June 1986.

[KARP 72] R.M. Karp, "Reducibility among combinatorial problems," in Complexity of Computer Computations, Plenum Press, pp.85-103, 1972.

[KLEI 75] L. Kleinrock and F.A. Tobagi, "Packet switching in radio channels: part I -- carrier sense multiple access modes and their throughput delay characteristics," IEEE Trans. Commun., vol.COM-23, pp.1400-1416, Dec. 1975.

[KLEI 87] L. Kleinrock and J.A. Silvester, "Spatial reuse in multihop packet radio networks," Proc. of IEEE, vol.75, No.1, pp.156-167, Jan. 1987.

[LEE 86] S.S. Lee and J.A. Silvester, "The effect of acknowledgements on the performance of distributed spread spectrum packet radio networks," Proc. of ICC'86, pp.1839-1846, June 1986.

[LEIN 87] B.M. Leiner, D.L.Nielson and F.A. Tobagi, "Issues in packet radio network design," Proc. of IEEE, vol.75, NO.1, pp.6-20, Jan. 1987.

[MCDI 79] C. McDiarmid, "Determining the chromatic number of a graph," SIAM J. Comput. Vol.8, No.1, pp.1-14, 1979.

[MITC 76] J. Mitchem, "On various algorithms for estimating the chromatic number of a graph," The Computer Journal, Vol.19, No.2, pp.182-183, 1976.

[PURS 87] M.B. Pursley, "The role of spread spectrum in packet radio networks," Proc. of IEEE, vol.75, No.1, pp.116-134, Jan. 1987.

[RAMA 89]R. Ramaswami and K. Parhi, "Distributed scheduling of broadcasts in a radio network," Proc. of INFOCOM'89, pp.497-504, April 1989.

[ROBE 75] L.G. Roberts, "ALOHA packet system with and without slots and capture," Comput. Commun. Rev., vol.5, pp.28-42, April 1975.

[SAST 84] A.R.K. Sastry, "Effect of acknowledgement traffic on the performance of slotted ALOHA-code division multiple access systems," IEEE Trans. Commun. vol.COM-32, No.11, pp.1219-1222, Nov. 1984.

[SOUS 88] E.S. Sousa and J.A. Silvester, "Spreading code protocols for distributed spread-spectrum packet radio networks," IEEE Trans. Commun. vol.COM-36, No.3, pp.272-281, March 1988.

[SOUZ 85] O. deSouza, M.S. Chen and R. Boorstyn, " A comparsion of the performance of protocols in packet radio networks," Proc. MILCOM'85, Boston, October 1985.

[TOBA 75] F.A. Tobagi and L. Kleinrock, "Packet switching in radio channels: part II -- the hidden terminal problem in carrier sense multiple access and the busy tone solution," IEEE Trans. Commun., COM-23, pp. 1417-1433, Dec. 1975.

[TOBA 80] F.A. Tobagi, "Multiaccess protocols in packet communications systems," IEEE Trans. Commun., vol. COM-28, No.4, pp.468-488, April 1980.

[TUCK 84] Alan Tucker, Applied Combinatorics, John Wiley & Sons, pp.66-79, 1984.